

OM protein - protein search, using sw model

Run on: August 24, 2004, 14:54:57 ; Search time 61.1194 Seconds
(without alignments)
69.343 Million cell updates/sec

Title: US-09-641-801-7
Perfect score: 81
Sequence: 1 VLEMKFPPPPQETVT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqpl980s:*
2: geneseqpl990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query Match	Length			
1	81	100.0	15	4	AAB72506	Aab72506 Colostrin
2	81	100.0	15	4	AAB59312	Aab59312 Ewe colos
3	81	100.0	15	4	AAB72252	Aab72252 Colostrin
4	81	100.0	15	4	AAB72538	Aab72538 Colostrin
5	81	100.0	15	5	AA014583	Aao14583 Neural ce
6	81	100.0	15	5	AAM51042	Aam51042 Colostrin
7	81	100.0	15	5	AAE20234	Aae20234 Colostrin
8	81	100.0	16	4	AAB59343	Aab59343 Ewe colos
9	51	63.0	164	3	AAG38918	Aag38918 Arabidops

10	51	63.0	218	3	AAG19665	Aag19665 Arabidops
11	51	63.0	239	3	AAG19664	Aag19664 Arabidops
12	51	63.0	268	3	AAG38917	Aag38917 Arabidops
13	51	63.0	289	3	AAG38916	Aag38916 Arabidops
14	51	63.0	289	4	AAE02472	Aae02472 Arabidops
15	51	63.0	289	6	ADB23136	Adb23136 Environme
16	51	63.0	289	7	ADE37165	Ade37165 Plant yie
17	51	63.0	319	7	ABO43111	Abo43111 A. thalia
18	51	63.0	319	7	ADC46623	Adc46623 Thalecres
19	51	63.0	319	7	ADD55790	Add55790 Thalecres
20	47	58.0	218	4	AAO06516	Aao06516 Human pol
21	46	56.8	141	4	ABG19150	Abg19150 Novel hum
22	45	55.6	78	4	AAM17645	Aaml17645 Peptide #
23	45	55.6	78	4	ABB36665	Abb36665 Peptide #
24	45	55.6	78	4	AAM30163	Aam30163 Peptide #
25	45	55.6	78	4	ABB31452	Abb31452 Peptide #
26	45	55.6	78	4	ABB22001	Abb22001 Protein #
27	45	55.6	78	4	AAM69824	Aam69824 Human bon
28	45	55.6	78	4	AAM57428	Aam57428 Human bra
29	45	55.6	78	4	ABG51515	Abg51515 Human liv
30	45	55.6	78	4	AAM05303	Aam05303 Peptide #
31	45	55.6	78	5	ABG39451	Abg39451 Human pep
32	45	55.6	114	4	ABB67485	Abb67485 Drosophil
33	45	55.6	366	4	ABG08125	Abg08125 Novel hum
34	44	54.3	166	3	AAG08618	Aag08618 Arabidops
35	44	54.3	166	7	ADD30054	Add30054 Plant yie
36	44	54.3	166	7	ADE31443	Ade31443 Plant yie
37	44	54.3	194	3	AAG08617	Aag08617 Arabidops
38	44	54.3	218	4	AAB63749	Aab63749 Human pro
39	44	54.3	289	5	ABB79218	Abb79218 Human pro
40	44	54.3	314	4	AAB63748	Aab63748 Human pro
41	44	54.3	329	4	AAM39525	Aam39525 Human pol
42	44	54.3	361	4	AAM14436	Aaml14436 Peptide #
43	44	54.3	361	4	AAM13889	Aaml13889 Peptide #
44	44	54.3	361	4	ABB32834	Abb32834 Peptide #
45	44	54.3	361	4	ABB33384	Abb33384 Peptide #

ALIGNMENTS

RESULT 1

AAB72506

ID AAB72506 standard; peptide; 15 AA.

XX

AC AAB72506;

XX

DT 09-MAY-2001 (first entry)

XX

DE Colostrinin peptide #7.

XX

KW Dermatological; oxidative stress regulator; colostrinin.

XX

OS Unidentified.

XX

PN WO200112650-A2.

XX

PD 22-FEB-2001.
 XX
 PF 17-AUG-2000; 2000WO-US022665.
 XX
 PR 17-AUG-1999; 99US-0149310P.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Stanton GJ, Hughes TK, Boldogh I;
 XX
 DR WPI; 2001-218342/22.
 XX
 PT Modulating oxidative stress level in a cell, involves contacting the cell
 PT with an oxidative stress regulator selected from colostrinin, its
 PT constituent peptide, analog or their combinations.
 XX
 PS Claim 6; Page 25; 48pp; English.
 XX
 CC The present invention relates to a method for modulating the oxidative
 CC stress level in a cell or a patient, comprising contacting the cell with,
 CC or administering to the patient, an oxidative stress regulator selected
 CC from colostrinin, or its constituent peptide (e.g. the present peptide),
 CC to change the level of an oxidising species in the cell. The method can
 CC be used to treat oxidative damage to skin, by decreasing or preventing an
 CC increase in the level of damage to a biomolecule of the patient
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 81; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 7.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
 |||||
 Db 1 VLEMKFPPPPQETVT 15

RESULT 2

AAB59312

ID AAB59312 standard; peptide; 15 AA.

XX

AC AAB59312;

XX

DT 21-MAR-2001 (first entry)

XX

DE Ewe colostrinin peptide fragment A-3.

XX

KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;

KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX

OS Ovis sp.

XX

PN WO200075173-A2.

XX

PD 14-DEC-2000.

XX

PF 02-JUN-2000; 2000WO-GB002128.

XX
 PR 02-JUN-1999; 99GB-00012852.
 XX
 PA (REGE-) REGEN THERAPEUTICS PLC.
 XX
 PI Georgiades JA;
 XX
 DR WPI; 2001-071058/08.
 XX
 PT Peptides having an N-terminal amino acid sequence isolated from
 PT colostrinin for treating e.g. disorders of the central nervous system and
 PT immune system, viral and bacterial infections, and diseases characterized
 PT by amyloid plaques.
 XX
 PS Claim 7; Page 27; 63pp; English.
 XX
 CC The present invention provides the sequences of a number of peptides
 CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide
 CC fragment of colostrum. These peptides can be used in the treatment of
 CC central nervous system disorders such as senile dementia, Parkinson's
 CC disease, Alzheimer's disease, psychosis and neurosis, immune system
 CC disorders such as bacterial and viral infections, to improve the
 CC development of a child's immune system, as a dietary supplement, and to
 CC promote the dissolution of beta-amyloid plaques
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 81; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 7.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
 |||||
 Db 1 VLEMKFPPPPQETVT 15

RESULT 3

AAB72252

ID AAB72252 standard; peptide; 15 AA.

XX

AC AAB72252;

XX

DT 14-MAY-2001 (first entry)

XX

DE Colostrinin derived cytokine inducing peptide SEQ ID 7.

XX

KW Colostrinin; immune response; cytokine; blood cell proliferation;
 KW central nervous system disorder; neurological disorder; mental disorder;
 KW dementia; neurodegenerative disease; Alzheimer's disease; psychosis;
 KW neurosis; infection.

XX

OS Synthetic.

XX

PN WO200111937-A2.

XX

PD 22-FEB-2001.

XX

PF 17-AUG-2000; 2000WO-US022818.
 XX
 PR 17-AUG-1999; 99US-0149311P.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (REGE-) REGEN THERAPEUTICS PLC.
 XX
 PI Stanton GJ, Hughes TK, Boldogh I, Georgiades J;
 XX
 DR WPI; 2001-202804/20.
 XX
 PT Inducing a cytokine and modulating an immune response, useful for
 PT treating central nervous system diseases and bacterial and viral
 PT infections, comprises administering colostrinin as an immunological
 PT regulator.
 XX
 PS Claim 1; Page 34; 50pp; English.
 XX
 CC Sequences AAB72246 - AAB72275 represent peptides derived from clostrinin,
 CC a proline rich polypeptide aggregate contained in colostrum. The peptides
 CC have immune response modulatory activity, and are capable of inducing
 CC cytokines. Colostrinin and its derived peptides are useful for inducing
 CC cytokine production, for modulating an immunological response and for
 CC inducing blood cell proliferation. The peptides are useful in the
 CC treatment of disorders of the central nervous system, neurological
 CC disorders, mental disorders, dementia, neurodegenerative diseases,
 CC Alzheimer's disease, motor neurone disease, psychosis, neurosis, chronic
 CC disorders of the immune system, bacterial and viral infections and
 CC acquired immunological deficiencies
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 81; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 7.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
 |||||
 Db 1 VLEMKFPPPPQETVT 15

RESULT 4

AAB72538

ID AAB72538 standard; peptide; 15 AA.

XX

AC AAB72538;

XX

DT 09-MAY-2001 (first entry)

XX

DE Colostrinin peptide #7.

XX

KW Neuroprotective; neural cell differentiation regulator; colostrinin;
 KW colostrum.

XX

OS Unidentified.

XX

PN WO200112651-A2.

XX
PD 22-FEB-2001.
XX
PF 17-AUG-2000; 2000WO-US022774.
XX
PR 17-AUG-1999; 99US-0149633P.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Boldogh I;
XX
DR WPI; 2001-226545/23.
XX
PT Use of colostrinin, its constituent peptide or analog as a neural cell
PT regulator, for promoting neural cell differentiation and treating damaged
PT neural cells in a patient.
XX
PS Claim 6; Page 21; 35pp; English.
XX
CC The present invention relates to a method for promoting neural cell
CC differentiation and treating damaged neural cells, using colostrinin and
CC colostrinin constituent peptides (e.g. the present peptide) as a neural
CC cell regulator. Colostrinin is a polypeptide complex found in colostrum
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 81; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 7.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
| | | | | | | | | | | | | | |
Db 1 VLEMKFPPPPQETVT 15

RESULT 5

AAO14583

ID AAO14583 standard; peptide; 15 AA.

XX

AC AAO14583;

XX

DT 27-MAY-2002 (first entry)

XX

DE Neural cell regulatory colostrinin peptide 7.

XX

KW Neural cell differentiation; neural cell regulator; colostrinin peptide;

KW neural cell formation; proline-rich polypeptide aggregate; colostrum;

KW neural cell treatment.

XX

OS Unidentified.

XX

FH Key Location/Qualifiers

FT Modified-site 15

FT /note= "Optional C-terminal amide"

XX

PN WO200213851-A1.

XX

PD 21-FEB-2002.
 XX
 PF 17-AUG-2000; 2000WO-US022777.
 XX
 PR 17-AUG-2000; 2000WO-US022777.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Boldogh I, Stanton JG, Hughes TK;
 XX
 DR WPI; 2002-269152/31.
 XX
 PT Promoting cell differentiation in a patient involves use of blood cell
 PT regulator selected from colostrinin, its constituent peptide and/or
 PT analog.
 XX
 PS Claim 7; Page 21; 37pp; English.
 XX
 CC The invention comprises a method for promoting cell differentiation (e.g.
 CC neural cell differentiation). The method involves contacting cells with a
 CC neural cell regulator (i.e. a colostrinin peptide) in order to change the
 CC cells in morphology to form neural cells. Colostrinin is a proline-rich
 CC polypeptide aggregate that is present in colostrum. The method of the
 CC invention is useful for promoting the differentiation of cells and for
 CC treating damaged neural cells in a patient. The present amino acid
 CC sequence represents a specifically claimed colostrinin peptide used in
 CC the method of the invention
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 81; DB 5; Length 15;
 Best Local Similarity 100.0%; Pred. No. 7.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
 |||||
 Db 1 VLEMKFPPPPQETVT 15

RESULT 6

AAM51042

ID AAM51042 standard; peptide; 15 AA.

XX

AC AAM51042;

XX

DT 30-MAY-2002 (first entry)

XX

DE Colostrinin constituent peptide.

XX

KW Colostrinin; colostrum; immunomodulator; cardiovascular;
 KW blood cell regulator; cytokine inducer; human.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Modified-site 15

FT /note= "optional C-terminal amidation"

XX
 PN WO200213849-A1.
 XX
 PD 21-FEB-2002.
 XX
 PF 17-AUG-2000; 2000WO-US022775.
 XX
 PR 17-AUG-2000; 2000WO-US022775.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (REGE-) REGEN THERAPEUTICS PLC.
 XX
 PI Stanton GJ, Hughes TK, Boldogh I, Georgiades J;
 XX
 DR WPI; 2002-269150/31.
 XX
 PT Modulation of blood cell proliferation in a patient involves use of blood
 PT cell regulator selected from colostrinin, its constituent peptide and/or
 PT analog.
 XX
 PS Claim 1; Page 34; 54pp; English.
 XX
 CC The present sequence is that of a colostrinin constituent peptide that is
 CC preferred for use as an immunological regulator and as a blood cell
 CC regulator in claimed methods of the invention. Methods are claimed for:
 CC inducing a cytokine in a cell by contact with an immunological regulator,
 CC where the cell is present in a cell culture, a tissue, an organ or an
 CC organism, and the cell is mammalian, including human; modulating an
 CC immune response in a cell by contact with the immunological regulator
 CC under conditions effective to induce a cytokine; modulating an immune
 CC response in a patient by administering an immunological regulator under
 CC conditions effective to induce a cytokine, where the immunological
 CC regulator is administered topically or as part of a dietary supplement,
 CC and where the immune response is specific or non specific, an interferon
 CC response or an antibody response; modulating blood cell proliferation by
 CC contacting blood cells with a blood cell regulator, where the blood cells
 CC are present in a cell culture or an organism, are mammalian or human, and
 CC where the blood cells are increased in number or differentiated; and a
 CC method for modulating blood cell proliferation in a patent. A claimed
 CC cytokine-inducing composition comprises a pharmaceutical carrier and an
 CC active agent such as the present peptide. Cytokines induced by this
 CC peptide in human leucocyte cultures include interferon-gamma, tumour
 CC necrosis factor-alpha, interleukin-4, interleukin-6 and interleukin-10
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 81; DB 5; Length 15;
 Best Local Similarity 100.0%; Pred. No. 7.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
 |||||
 Db 1 VLEMKFPPPPQETVT 15

RESULT 7
 AAE20234

ID AAE20234 standard; peptide; 15 AA.
XX
AC AAE20234;
XX
DT 18-JUN-2002 (first entry)
XX
DE Colostrinin constituent peptide #7.
XX
KW Blood cell regulator; colostrinin; constituent peptide; oxidative stress;
KW therapy; oxidative damage; skin; aging; wound healing; cell replacement;
KW tissue; organ; cosmetic procedure; repair; regeneration; preservation;
KW transplantation; implantation; dermatological; vulnerary.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 15
FT /note= "Optionally C-terminal amide"
XX
PN WO200213850-A1.
XX
PD 21-FEB-2002.
XX
PF 17-AUG-2000; 2000WO-US022776.
XX
PR 17-AUG-2000; 2000WO-US022776.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Stanton GJ, Hughes TK, Boldogh I;
XX
DR WPI; 2002-269151/31.
XX
PT Composition useful for the modulation of blood cell proliferation in a
PT patient comprises a blood cell regulator selected from colostrinin, its
PT constituent peptide and/or analog.
XX
PS Claim 6; Page 25; 51pp; English.
XX
CC The invention relates to a composition which comprises a blood cell
CC regulator selected from colostrinin, its constituent peptide and/or
CC analogue. The invention is used for modulating the oxidative stress level
CC in a cell e.g. mammalian or human cell present in a cell culture, tissue,
CC organ, or organism; or for treating oxidative damage to the skin of a
CC patient e.g. animal or human; to modulate oxidative stress during/ after
CC a premature birth or normal birth, preventing/delaying aging in a
CC patient, enhancing wound healing, and the reduction of side effects of
CC cosmetic procedures. The method changes the level of an oxidising species
CC in the cell, such as decreases or prevents increase in the level of
CC damage to a biomolecule of the patient selected from DNA, protein and/or
CC lipid, compared to the same conditions when the oxidative stress
CC regulator is not present. The modulation of oxidative stress results in
CC enhanced repair, regeneration, and replacement of cells, tissues and
CC organs (e.g. kidney, liver, pancreas, skin, and the other internal and
CC external organs), as well as enhanced preservation of such organs for
CC transplantation, implantation, or scientific research. The present
CC sequence is a colostrinin constituent peptide

XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 81; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 7.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
| | | | | | | | | | | | | | |
Db 1 VLEMKFPPPPQETVT 15

RESULT 8

AAB59343

ID AAB59343 standard; peptide; 16 AA.

XX

AC AAB59343;

XX

DT 21-MAR-2001 (first entry)

XX

DE Ewe colostrinin peptide fragment derived sequence #3.

XX

KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;

KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX

OS Ovis sp.

XX

PN WO200075173-A2.

XX

PD 14-DEC-2000.

XX

PF 02-JUN-2000; 2000WO-GB002128.

XX

PR 02-JUN-1999; 99GB-00012852.

XX

PA (REGE-) REGEN THERAPEUTICS PLC.

XX

PI Georgiades JA;

XX

DR WPI; 2001-071058/08.

XX

PT Peptides having an N-terminal amino acid sequence isolated from
PT colostrinin for treating e.g. disorders of the central nervous system and
PT immune system, viral and bacterial infections, and diseases characterized
PT by amyloid plaques.

XX

PS Claim 8; Page 27; 63pp; English.

XX

CC The present invention provides the sequences of a number of peptides
CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide
CC fragment of colostrum. These peptides can be used in the treatment of
CC central nervous system disorders such as senile dementia, Parkinson's
CC disease, Alzheimer's disease, psychosis and neurosis, immune system
CC disorders such as bacterial and viral infections, to improve the
CC development of a child's immune system, as a dietary supplement, and to
CC promote the dissolution of beta-amyloid plaques

XX

SQ Sequence 16 AA;

Query Match 100.0%; Score 81; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 7.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
|||||
Db 2 VLEMKFPPPPQETVT 16

RESULT 9

AAG38918

ID AAG38918 standard; protein; 164 AA.

XX

AC AAG38918;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 48081.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.

PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.

PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 63.0%; Score 51; DB 3; Length 164;
 Best Local Similarity 69.2%; Pred. No. 13;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
 | | | | | | | | | |
 Db 59 VTEMVMPPPPQQT 71

RESULT 10
 AAG19665
 ID AAG19665 standard; protein; 218 AA.
 XX
 AC AAG19665;

XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 21552.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.

PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.

PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.

PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 63.0%; Score 51; DB 3; Length 218;
 Best Local Similarity 69.2%; Pred. No. 17;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
 | || |||||:
 Db 163 VTEMVMPPPPQQT 175

RESULT 11

AAG19664

ID AAG19664 standard; protein; 239 AA.

XX

AC AAG19664;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 21551.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR	25-FEB-1999;	99US-0121825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.

PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.

PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.

PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 63.0%; Score 51; DB 3; Length 239;
 Best Local Similarity 69.2%; Pred. No. 19;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
 | || |||||:
 Db 184 VTEMVMPPPPQQT 196

RESULT 12

AAG38917

ID AAG38917 standard; protein; 268 AA.

XX

AC AAG38917;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 48080.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.

PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.

PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.

PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 63.0%; Score 51; DB 3; Length 268;
 Best Local Similarity 69.2%; Pred. No. 21;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLEMKFPPPPQET 13

Db 163 VTEMVMPPPPQQT 175

RESULT 13

AAG38916

ID AAG38916 standard; protein; 289 AA.
XX
AC AAG38916;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48079.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.

PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.

PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.

PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 63.0%; Score 51; DB 3; Length 289;
 Best Local Similarity 69.2%; Pred. No. 23;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
 | || |||||:|
 Db 184 VTEMVMPPPPQQT 196

RESULT 14

AAE02472

ID AAE02472 standard; protein; 289 AA.

XX

AC AAE02472;

XX

DT 10-AUG-2001 (first entry)

XX

DE Arabidopsis thaliana transcription factor G502.

XX

KW Transcription factor; environmental stress tolerance; gene therapy;
 KW plant structure; plant development.

XX
 OS Arabidopsis thaliana.
 XX
 FH Key Location/Qualifiers
 FT Domain 10. .155
 FT /note= "Conserved domain"
 XX
 PN WO200136598-A1.
 XX
 PD 25-MAY-2001.
 XX
 PF 14-NOV-2000; 2000WO-US031458.
 XX
 PR 17-NOV-1999; 99US-0166228P.
 PR 17-APR-2000; 2000US-0197899P.
 PR 22-AUG-2000; 2000US-0227439P.
 XX
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.
 PA (PINE/) PINEDA O.
 PA (YUGG/) YU G.
 PA (CREE/) CREELMAN R.
 PA (RIEC/) RIECHMANN J L.
 PA (HEAR/) HEARD J.
 PA (RATC/) RATCLIFFE O.
 PA (REUB/) REUBER L.
 PA (KEDD/) KEDDIE J.
 XX
 PI Pineda O, Yu G, Creelman R, Riechmann JL, Heard J, Ratcliffe O;
 PI Reuber L, Keddie J;
 XX
 DR WPI; 2001-336000/35.
 DR N-PSDB; AAD06452.
 XX
 PT Nucleic acids encoding plant transcription factor polypeptides, useful
 PT for altering the environmental stress tolerance characteristics of
 PT plants.
 XX
 PS Claim 4; Page 73-74; 116pp; English.
 XX
 CC The present amino acid sequence is Arabidopsis thaliana transcription
 CC factor. This novel transcription factor is useful for modifying a plant's
 CC phenotype in desirable ways, such as modifying a plants environmental
 CC stress. The transcription factor is encoded by environmental stress
 CC tolerance gene derived from Arabidopsis thaliana. The transcription
 CC factors and the genes encoding them are used to alter the structure and
 CC developmental characteristics of plants such as soybean, wheat, corn,
 CC potato, cotton, rice, oilseed rape, sunflower, alfalfa, sugarcane, turf,
 CC banana, blackberry, blueberry, strawberry, raspberry, cantaloupe, carrot,
 CC cauliflower, coffee, onion, cucumber, eggplant, grapes, honey dew,
 CC lettuce, mango, melon, papaya, peas, peppers, pineapple, spinach, squash,
 CC sweet corn, tobacco, tomato, watermelon, rosaceous fruits and/or
 CC vegetable brassicas. These sequences are also used for modifying traits
 CC associated with environmental stress tolerance, such as freezing,
 CC chilling, heat, drought, water saturation, salt, photoconditions,
 CC radiation and ozone. The transcription factors are used in gene therapy
 XX
 SQ Sequence 289 AA;

Query Match 63.0%; Score 51; DB 4; Length 289;
Best Local Similarity 69.2%; Pred. No. 23;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
| | | | | : |
Db 184 VTEMVMPPPPQQT 196

RESULT 15

ADB23136

ID ADB23136 standard; protein; 289 AA.
XX
AC ADB23136;
XX
DT 20-NOV-2003 (first entry)
XX
DE Environmental stress-responsive promoter-related protein, SEQ ID 140.
XX
KW Plant; environmental stress; promoter.
XX
OS Arabidopsis thaliana.
XX
PN WO2003044190-A1.
XX
PD 30-MAY-2003.
XX
PF 15-NOV-2002; 2002WO-JP011955.
XX
PR 19-NOV-2001; 2001JP-00353038.
PR 29-JAN-2002; 2002JP-00020329.
XX
PA (RIKE) RIKEN KK.
XX
PI Shinozaki K, Seki M, Fujita M;
XX
DR WPI; 2003-468771/44.
DR N-PSDB; ADB23135.
XX
PT Environmental stress-responsive promoters, useful in constructing and
PT breeding stress-resistant plants for applications in agriculture.
XX
PS Claim 9; Page 432-434; 520pp; Japanese.
XX
CC The present invention relates to novel environmental stress-responsive
CC promoters (ADB23163-ADB23252) from Arabidopsis thaliana. The promoters
CC are useful in constructing and breeding stress-resistant plants for
CC applications in agriculture and horticulture. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 289 AA;

Query Match 63.0%; Score 51; DB 6; Length 289;
Best Local Similarity 69.2%; Pred. No. 23;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
 | | | | | | | : |
Db 184 VTEMVMPPPPQQT 196

Search completed: August 24, 2004, 15:42:35
Job time : 65.1194 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:33:13 ; Search time 16.4552 Seconds
 (without alignments)
 47.060 Million cell updates/sec

Title: US-09-641-801-7
 Perfect score: 81
 Sequence: 1 VLEMKFPPPPQETVT 15

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Issued_Patents_AA:*
 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	81	100.0	15	4	US-09-641-803-7	Sequence 7, Appli
2	51	63.0	319	4	US-09-533-029-38	Sequence 38, Appl
3	46	56.8	300	4	US-09-252-991A-29412	Sequence 29412, A
4	44	54.3	521	2	US-08-933-227-1	Sequence 1, Appli
5	43	53.1	521	2	US-08-933-227-3	Sequence 3, Appli
6	42	51.9	424	4	US-09-252-991A-20695	Sequence 20695, A
7	41	50.6	130	4	US-09-328-352-6607	Sequence 6607, Ap
8	41	50.6	1250	1	US-08-441-139-9	Sequence 9, Appli
9	40	49.4	82	4	US-09-621-976-4586	Sequence 4586, Ap
10	40	49.4	91	4	US-09-621-976-4585	Sequence 4585, Ap
11	40	49.4	366	3	US-09-176-657-2	Sequence 2, Appli

12	40	49.4	366	4	US-09-421-299-2	Sequence 2, Appli
13	39	48.1	23	1	US-08-268-251-56	Sequence 56, Appl
14	39	48.1	23	5	PCT-US93-01112-56	Sequence 56, Appl
15	39	48.1	153	4	US-09-252-991A-31659	Sequence 31659, A
16	39	48.1	450	4	US-09-252-991A-26266	Sequence 26266, A
17	39	48.1	995	5	PCT-US95-04910-14	Sequence 14, Appl
18	38	46.9	102	4	US-09-198-452A-475	Sequence 475, App
19	38	46.9	181	4	US-09-252-991A-23085	Sequence 23085, A
20	38	46.9	345	5	PCT-US95-03323A-2	Sequence 2, Appli
21	38	46.9	460	4	US-09-056-556-184	Sequence 184, App
22	38	46.9	460	4	US-09-072-596-179	Sequence 179, App
23	38	46.9	460	4	US-09-072-967-184	Sequence 184, App
24	38	46.9	469	4	US-09-489-039A-13565	Sequence 13565, A
25	38	46.9	475	4	US-09-370-838-193	Sequence 193, App
26	38	46.9	570	4	US-09-489-039A-7328	Sequence 7328, Ap
27	38	46.9	913	1	US-08-445-640-4	Sequence 4, Appli
28	38	46.9	913	3	US-08-170-558-4	Sequence 4, Appli
29	38	46.9	913	3	US-08-447-314-4	Sequence 4, Appli
30	38	46.9	913	3	US-08-445-461-4	Sequence 4, Appli
31	38	46.9	913	4	US-09-140-378A-2	Sequence 2, Appli
32	38	46.9	919	1	US-08-336-343A-2	Sequence 2, Appli
33	38	46.9	1052	2	US-08-852-806-2	Sequence 2, Appli
34	38	46.9	1052	3	US-09-163-669-2	Sequence 2, Appli
35	38	46.9	1105	3	US-08-999-774A-2	Sequence 2, Appli
36	38	46.9	1178	4	US-09-489-039A-9944	Sequence 9944, Ap
37	38	46.9	1912	4	US-09-495-714C-2	Sequence 2, Appli
38	38	46.9	1977	4	US-09-495-714C-4	Sequence 4, Appli
39	38	46.9	1985	4	US-09-495-714C-6	Sequence 6, Appli
40	38	46.9	2710	2	US-08-568-459A-12	Sequence 12, Appl
41	38	46.9	2710	2	US-08-487-826B-12	Sequence 12, Appl
42	38	46.9	2710	4	US-09-210-288-12	Sequence 12, Appl
43	38	46.9	3060	2	US-08-487-826B-14	Sequence 14, Appl
44	37.5	46.3	65	4	US-09-252-991A-20215	Sequence 20215, A
45	37	45.7	18	3	US-08-630-916A-109	Sequence 109, App

ALIGNMENTS

RESULT 1

US-09-641-803-7

; Sequence 7, Application US/09641803

; Patent No. 6500798

; GENERAL INFORMATION:

; APPLICANT: STANTON, G. John

; APPLICANT: HUGHES, Thomas K.

; APPLICANT: BOLDOGH, Istvan

; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND

; TITLE OF INVENTION: ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS

; FILE REFERENCE: 265.00220101

; CURRENT APPLICATION NUMBER: US/09/641,803

; CURRENT FILING DATE: 2000-08-17

; PRIOR APPLICATION NUMBER: 60/149,310

; PRIOR FILING DATE: 1999-08-17

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 7

```
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-09-641-803-7
```

```
Query Match          100.0%; Score 81; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 VLEMKFPPPPQETVT 15
        |||||
Db      1 VLEMKFPPPPQETVT 15
```

RESULT 2

```
US-09-533-029-38
; Sequence 38, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502
US-09-533-029-38
```

```
Query Match          63.0%; Score 51; DB 4; Length 319;
Best Local Similarity 69.2%; Pred. No. 2.4;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy      1 VLEMKFPPPPQET 13
        ||| |||||
```

Db 184 VTEMVMPPPPQQT 196

RESULT 3

US-09-252-991A-29412
; Sequence 29412, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29412
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29412

Query Match 56.8%; Score 46; DB 4; Length 300;
Best Local Similarity 80.0%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPP 10
|:||||| |
Db 221 VVEMKFPPDP 230

RESULT 4

US-08-933-227-1
; Sequence 1, Application US/08933227
; Patent No. 5965394
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Guegler, Karl
; APPLICANT: Corley, Neil
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN IMPORTIN ALPHA HOMOLOG
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

```

;   SOFTWARE:  FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/933,227
;     FILING DATE:  Filed Herewith
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Billings, Lucy J.
;     REGISTRATION NUMBER:  36,749
;     REFERENCE/DOCKET NUMBER:  PF-0394 US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  650-855-0555
;     TELEFAX:  650-845-4166
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  521 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       LIBRARY:  THP1PLB01
;       CLONE:  10403
US-08-933-227-1

```

```

Query Match          54.3%;  Score 44;  DB 2;  Length 521;
Best Local Similarity 80.0%;  Pred. No. 45;
Matches      8;  Conservative      0;  Mismatches      2;  Indels      0;  Gaps      0;

```

```

Qy          5 KFPPPPQETV 14
             | |||| |||
Db          231 KDPPPPMETV 240

```

RESULT 5

US-08-933-227-3

```

; Sequence 3, Application US/08933227
; Patent No. 5965394
;   GENERAL INFORMATION:
;     APPLICANT:  Bandman, Olga
;     APPLICANT:  Guegler, Karl
;     APPLICANT:  Corley, Neil
;     APPLICANT:  Shah, Purvi
;     TITLE OF INVENTION:  HUMAN IMPORTIN ALPHA HOMOLOG
;     NUMBER OF SEQUENCES:  5
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;       STREET:  3174 Porter Dr.
;       CITY:  Palo Alto
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94304
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Diskette
;     COMPUTER:  IBM Compatible
;     OPERATING SYSTEM:  DOS
;     SOFTWARE:  FastSEQ for Windows Version 2.0

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,227
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0394 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1928975
US-08-933-227-3

```

```

Query Match          53.1%; Score 43; DB 2; Length 521;
Best Local Similarity 70.0%; Pred. No. 64;
Matches      7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      5 KFPPPPQETV 14
        | ||| ||:
Db      231 KDPPPPMETI 240

```

RESULT 6

```

US-09-252-991A-20695
; Sequence 20695, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20695
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20695

```

```

Query Match          51.9%; Score 42; DB 4; Length 424;

```

Best Local Similarity 77.8%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 PPPPQETVT 15
||||: |||
Db 34 PPPRPPTVT 42

RESULT 7

US-09-328-352-6607
; Sequence 6607, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6607
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6607

Query Match 50.6%; Score 41; DB 4; Length 130;
Best Local Similarity 64.3%; Pred. No. 30;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETV 14
|| |:|| ||| |
Db 35 VLIMRFPRSPQEWV 48

RESULT 8

US-08-441-139-9
; Sequence 9, Application US/08441139
; Patent No. 5773245
; GENERAL INFORMATION:
; APPLICANT: Wittrup, Dr. Karl D.
; APPLICANT: Robinson, Anne S.
; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
; TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

;      SOFTWARE: PatentIn Release #1.0, Version #1.25
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/441,139
;      FILING DATE: 15-MAY-1995
;      CLASSIFICATION: 435
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US 08/089,997
;      FILING DATE: 06-JUL-1993
;      ATTORNEY/AGENT INFORMATION:
;      NAME: DiGiglio, Frank S.
;      REGISTRATION NUMBER: 31,346
;      REFERENCE/DOCKET NUMBER: 8646
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 516-742-4343
;      TELEFAX: 516-742-4366
;      TELEX: 230 901 SANS UR
;      INFORMATION FOR SEQ ID NO: 9:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 1250 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-08-441-139-9

```

```

Query Match          50.6%; Score 41; DB 1; Length 1250;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches      6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      3 EMKFPPPP 10
        : |||||
Db      210 DFKFPPPP 217

```

```

RESULT 9
US-09-621-976-4586
; Sequence 4586, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4586
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4586

```

```

Query Match          49.4%; Score 40; DB 4; Length 82;
Best Local Similarity 66.7%; Pred. No. 26;
Matches      6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```


Qy 6 FPPPPQETV 14
| | | | | :
Db 26 FPPPPQNGI 34

RESULT 10

US-09-621-976-4585
; Sequence 4585, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4585
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4585

Query Match 49.4%; Score 40; DB 4; Length 91;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 FPPPPQETV 14
| | | | | :
Db 10 FPPPPQNGI 18

RESULT 11

US-09-176-657-2
; Sequence 2, Application US/09176657
; Patent No. 6020164
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Lu, Aina
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: HUMAN RNA BINDING PROTEINS
; FILE REFERENCE: PF-0611 US
; CURRENT APPLICATION NUMBER: US/09/176,657
; CURRENT FILING DATE: 1998-10-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 1250374

US-09-176-657-2

Query Match 49.4%; Score 40; DB 3; Length 366;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 FPPPPQETV 14
| | | | | :
Db 30 FPPPPQNGI 38

RESULT 12

US-09-421-299-2

; Sequence 2, Application US/09421299
; Patent No. 6524579
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Lu, Aina
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: HUMAN RNA BINDING PROTEINS
; FILE REFERENCE: PF-0611 US
; CURRENT APPLICATION NUMBER: US/09/421,299
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: 09/176,657
; EARLIER FILING DATE: 1998-10-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 1250374
US-09-421-299-2

Query Match 49.4%; Score 40; DB 4; Length 366;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 FPPPPQETV 14
| | | | | :
Db 30 FPPPPQNGI 38

RESULT 13

US-08-268-251-56

; Sequence 56, Application US/08268251
; Patent No. 5585475
; GENERAL INFORMATION:
; APPLICANT: Jamieson, Gordon A
; APPLICANT: Dedman, John R
; APPLICANT: Kaetzel, Marcia A
; TITLE OF INVENTION: Calmodulin-Binding Peptides
; NUMBER OF SEQUENCES: 59

```

;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Chiron Corporation
;   STREET:    4560 Horton Street
;   CITY:      Emeryville
;   STATE:     CA
;   COUNTRY:   USA
;   ZIP:       94608
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:   IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:   WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/268,251
;   FILING DATE:
;   CLASSIFICATION:  514
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  US 07/831,219
;   FILING DATE:    06-FEB-1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME:   Green, Grant D
;   REGISTRATION NUMBER:  31,259
;   REFERENCE/DOCKET NUMBER:  272.001
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (510) 601-2706
;   TELEFAX:   (510) 655-3542
;   INFORMATION FOR SEQ ID NO:  56:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  23 amino acids
;   TYPE:    amino acid
;   STRANDEDNESS:  single
;   TOPOLOGY:  linear
;   MOLECULE TYPE:  peptide
US-08-268-251-56

```

```

Query Match          48.1%;  Score 39;  DB 1;  Length 23;
Best Local Similarity 85.7%;  Pred. No. 9.8;
Matches      6;  Conservative  0;  Mismatches   1;  Indels    0;  Gaps    0;

```

```

Qy          4 MKFP PPP 10
             | ||||
Db          15 MNFP PPP 21

```

RESULT 14

PCT-US93-01112-56

; Sequence 56, Application PC/TUS9301112

; GENERAL INFORMATION:

```

;   APPLICANT:  Jamieson, Gordon A
;   APPLICANT:  Dedman, John R
;   APPLICANT:  Kaetzel, Marcia A
;   TITLE OF INVENTION:  Calmodulin-Binding Peptides
;   NUMBER OF SEQUENCES:  59
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Chiron Corporation
;   STREET:    4560 Horton Street
;   CITY:      Emeryville

```

```

; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01112
; FILING DATE: 19930208
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,219
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 272.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-01112-56

```

```

Query Match          48.1%; Score 39; DB 5; Length 23;
Best Local Similarity 85.7%; Pred. No. 9.8;
Matches      6; Conservative      0; Mismatches      1; Indels      0; Gaps      0;

```

```

Qy      4 MKFPPPP 10
        | |||||
Db      15 MNFPPPP 21

```

RESULT 15

```

US-09-252-991A-31659
; Sequence 31659, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

```

; SEQ ID NO 31659
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31659

Query Match 48.1%; Score 39; DB 4; Length 153;
Best Local Similarity 66.7%; Pred. No. 71;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEMKFPPPP 10
:| |||||
Db 80 VESSFPPPP 88

Search completed: August 24, 2004, 15:55:18
Job time : 17.4552 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:26:28 ; Search time 14.5522 Seconds
(without alignments)
99.151 Million cell updates/sec

Title: US-09-641-801-7
Perfect score: 81
Sequence: 1 VLEMKFPPPPQETVT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result			%				ID	Description
	No.	Score	Match	Length	DB	Query		
1	51	63.0	229	2	S37101	ATAF1 protein - Ar		
2	51	63.0	289	2	E86148	T1N6.12 protein -		
3	49	60.5	180	2	E84768	hypothetical prote		
4	46	56.8	148	2	F83542	hypothetical prote		
5	46	56.8	272	2	S24375	hypothetical prote		
6	44	54.3	204	2	B86410	protein F3M18.20 [
7	44	54.3	528	2	T52092	DNA-binding protei		
8	44	54.3	1203	2	S26650	DNA-binding protei		
9	43	53.1	196	2	T29602	hypothetical prote		
10	43	53.1	330	2	S37595	mucin JUL10 - huma		
11	43	53.1	521	2	JC5505	DNA helicase Q1 in		
12	43	53.1	543	2	S35047	mucin JUL7 - human		
13	43	53.1	599	2	T39990	hypothetical prote		

14	43	53.1	610	2	S71110	abcl protein homol
15	43	53.1	688	2	T09941	transcription fact
16	43	53.1	709	2	T26874	hypothetical prote
17	42	51.9	79	2	E87607	hypothetical prote
18	42	51.9	153	2	C86328	protein F18014.31
19	42	51.9	181	2	T22884	hypothetical prote
20	42	51.9	182	2	A46436	myofibrillar prote
21	42	51.9	291	2	A30558	B-cell surface ant
22	42	51.9	440	2	D86328	protein F18014.33
23	42	51.9	448	2	B87635	transcription regu
24	42	51.9	475	2	JX0344	legumain (EC 3.4.2
25	42	51.9	811	2	E72003	hypothetical prote
26	42	51.9	913	2	D71463	probable ATP-depen
27	41	50.6	181	2	S44765	C29E4.9 protein -
28	41	50.6	213	2	D20647	heat shock protein
29	41	50.6	270	2	A84549	probable C2H2-type
30	41	50.6	287	2	T01919	probable Ap2 domai
31	41	50.6	288	2	T45715	hypothetical prote
32	41	50.6	475	2	T36921	hypothetical prote
33	41	50.6	493	2	T12044	probable legumain
34	41	50.6	920	2	F81725	cell division prot
35	41	50.6	1041	2	T15521	hypothetical prote
36	41	50.6	1196	2	T23832	protein-tyrosine k
37	41	50.6	1250	2	A39578	SSD1 protein - yea
38	41	50.6	2845	2	I49505	adenomatous polypo
39	40.5	50.0	1051	2	A39712	kinase-like protei
40	40.5	50.0	1070	2	JC4593	protein-tyrosine k
41	40	49.4	195	2	T49012	calmodulin-like pr
42	40	49.4	205	2	B75253	DedA family protei
43	40	49.4	256	1	A48103	ribosomal protein
44	40	49.4	308	2	C30554	Ig heavy chain C r
45	40	49.4	339	2	A84095	transcription regu

ALIGNMENTS

RESULT 1

S37101

ATAF1 protein - Arabidopsis thaliana (fragment)

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Sep-1997

C;Accession: S37101

R;Rueth, J.; Schweyen, R.; Hirt, H.

submitted to the EMBL Data Library, August 1993

A;Reference number: S37100

A;Accession: S37101

A;Molecule type: mRNA

A;Residues: 1-229 <RUE>

A;Cross-references: EMBL:X74755; NID:g398603; PID:e85756; PID:g1345506

C;Genetics:

A;Gene: ATAF1

Query Match 63.0%; Score 51; DB 2; Length 229;

Best Local Similarity 69.2%; Pred. No. 1.5;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
| || |||||:|
Db 124 VTEMVMPPPPQQT 136

RESULT 2

E86148

T1N6.12 protein - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Nov-2001

C;Accession: E86148

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: E86148

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-289 <STO>

A;Cross-references: GB:AE005172; NID:g8671840; PIDN:AAF78403.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

Query Match 63.0%; Score 51; DB 2; Length 289;
Best Local Similarity 69.2%; Pred. No. 2;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
| || |||||:|
Db 184 VTEMVMPPPPQQT 196

RESULT 3

E84768

hypothetical protein At2g35430 [imported] - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C;Accession: E84768

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams,

M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: E84768

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-180 <STO>

A;Cross-references: GB:AE002093; NID:g3608145; PIDN:AAC36178.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g35430

A;Map position: 2

Query Match 60.5%; Score 49; DB 2; Length 180;
Best Local Similarity 58.8%; Pred. No. 2.3;
Matches 10; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

Qy 3 EMKFPPPP---QETVT 15
|:: |||| ||||
Db 100 ELRLPPPPPNWQETVT 116

RESULT 4

F83542

hypothetical protein PA0822 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C;Species: *Pseudomonas aeruginosa*

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C;Accession: F83542

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brinkman, F.S.L.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber, R.L.; Goltry, L.; Tolentino, E.; Westbrook-Wadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, R.M.; Smith, K.A.; Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.; Hancock, R.E.W.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen.

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: F83542

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-148 <STO>

A;Cross-references: GB:AE004517; GB:AE004091; NID:g9946710; PIDN:AAG04211.1;

GSPDB:GN00131; PASP:PA0822

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA0822

Query Match 56.8%; Score 46; DB 2; Length 148;
Best Local Similarity 80.0%; Pred. No. 5.3;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPP 10
|:||||| |

RESULT 5

S24375

hypothetical protein 3 - *Pseudomonas* sp. (DSM 6898) plasmid pKB740C;Species: *Pseudomonas* sp.

C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 08-Oct-1999

C;Accession: S24375

R;Altenschmidt, U.; Bokranz, M.; Fuchs, G.

Eur. J. Biochem. 207, 715-722, 1992

A;Title: Novel aerobic 2-aminobenzoate metabolism. Nucleotide sequence of the plasmid carrying the gene for the flavoprotein 2-aminobenzoyl-CoA monooxygenase/reductase in a denitrifying *Pseudomonas* sp.

A;Reference number: S24374; MUID:92339462; PMID:1633822

A;Accession: S24375

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-272 <ALT>

A;Cross-references: EMBL:X66604; NID:g45865; PIDN:CAA47169.1; PID:g45867

C;Genetics:

A;Genome: plasmid

Query Match	56.8%;	Score 46;	DB 2;	Length 272;
Best Local Similarity	58.3%;	Pred. No. 10;		
Matches	7;	Conservative	2;	Mismatches 3; Indels 0; Gaps 0;

Qy	2 LEMKFPPPPQET 13
	: :
Db	153 LRRRFPPPPDDT 164

RESULT 6

B86410

protein F3M18.20 [imported] - *Arabidopsis thaliana*C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: B86410

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.;

Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.;

Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.;

Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.;

Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.;

Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.;

Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-

Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu,

S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda,

M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.;

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.;

Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken,

S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.;

Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: B86410
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-204 <STO>
A;Cross-references: GB:AE005172; NID:g6560756; PIDN:AAF16756.1; GSPDB:GN00141
C;Genetics:
A;Gene: F3M18.20
A;Map position: 1

Query Match 54.3%; Score 44; DB 2; Length 204;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LEMKFPPPPQ 11
|:: |||||:
Db 194 LDLNFPPPPPE 203

RESULT 7

T52092

DNA-binding protein WRKY2 [imported] - common tobacco

C;Species: Nicotiana tabacum (common tobacco)

C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000

C;Accession: T52092

R;Wang, Z.; Yang, P.; Fan, B.; Chen, Z.

Plant J. 16, 515-522, 1998

A;Title: An oligo selection procedure for identification of sequence-specific DNA-binding activities associated with the plant defence response.

A;Reference number: Z25947; MUID:99097845; PMID:9881170

A;Accession: T52092

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-528 <WAN>

A;Cross-references: EMBL:AF096299; PIDN:AAD16139.1

C;Genetics:

A;Gene: WRKY2

Query Match 54.3%; Score 44; DB 2; Length 528;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 FPPPPQETVT 15
|||||: ||:
Db 79 FPPPPPEMTVS 88

RESULT 8

S26650

DNA-binding protein 5 - human

C;Species: Homo sapiens (man)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000

C;Accession: S26650; PN0099; I57397

R;Mattioni, T.; Hume, C.R.; Konigorski, S.; Hayes, P.; Osterweil, Z.; Lee, J.S.
Chromosoma 101, 618-624, 1992

A;Title: A cDNA clone for a novel nuclear protein with DNA binding activity.

A;Reference number: S26650; MUID:93048367; PMID:1424986

A;Accession: S26650

A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-1203 <MAT>
 A;Cross-references: EMBL:X63071
 R;Berdichevsky, F.B.; Chumakov, I.M.; Kisselev, L.L.
 Mol. Biol. (Mosk.) 22, 794-801, 1988
 A;Title: Determination of the nucleotide sequence of the human genome son3
 fragment: identification of a new protein bearing an unusual structure and
 homology with DNA binding proteins.
 A;Reference number: PN0099; MUID:89039788; PMID:3054499
 A;Accession: PN0099
 A;Molecule type: mRNA
 A;Residues: 'R',574-699,'PH',702-1055 <BER>
 A;Experimental source: embryo
 A;Note: this paper is in Russian, with an English abstract
 R;Berdichevsky, F.B.; Chumakov, I.M.; Kiselev, L.L.
 Mol. Biol. 22, 639-646, 1988
 A;Title: Determination of the nucleotide sequence of the son3 fragment of the
 human genome: Identification of a new protein with an unusual structure and
 homology with DNA-binding proteins.
 A;Reference number: I57397
 A;Accession: I57397
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 'R',574-699,'PH',702-1055 <RES>
 A;Cross-references: GB:M36428; NID:g338289; PIDN:AAA36624.1; PID:g338290
 C;Genetics:
 A;Introns: 790/3
 C;Keywords: DNA binding
 F;741-928/Region: arginine/serine-rich repeats

Query Match 54.3%; Score 44; DB 2; Length 1203;
 Best Local Similarity 58.3%; Pred. No. 1.1e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EMKFPPPPQETV 14
 | : ||||:|:
 Db 585 EKEVPPPPKETL 596

RESULT 9
 T29602
 hypothetical protein ZK816.4 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
 C;Accession: T29602
 R;Nhan, M.; Le, T.
 submitted to the EMBL Data Library, November 1995
 A;Description: The sequence of C. elegans cosmid ZK816.
 A;Reference number: Z20649
 A;Accession: T29602
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-196 <NHA>
 A;Cross-references: EMBL:U41018; PIDN:AAA82325.1; CESP:ZK816.4
 C;Genetics:
 A;Gene: CESP:ZK816.4

A;Introns: 40/3; 66/2; 105/2; 148/1

C;Superfamily: Caenorhabditis elegans hypothetical protein ZK816.4

Query Match 53.1%; Score 43; DB 2; Length 196;
Best Local Similarity 70.0%; Pred. No. 21;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EMKFPPPPQE 12
: ||||| ||
Db 19 QFKFPPPTQE 28

RESULT 10

S37595

mucin JUL10 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Oct-1999

C;Accession: S37595; S35048

R;Aubert, J.

submitted to the EMBL Data Library, September 1993

A;Reference number: S37593

A;Accession: S37595

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-330 <AUB>

A;Cross-references: EMBL:X74956; NID:g407052; PIDN:CAA52911.1; PID:g407053

R;Dufosse, J.; Porchet, N.; Audie, J.P.; Guyonnet Duperat, V.; Laine, A.; van-Seuning, I.; Marrakchi, S.; Degand, P.; Aubert, J.P.

Biochem. J. 293, 329-337, 1993

A;Title: Degenerate 87-base-pair tandem repeats create hydrophilic/hydrophobic alternating domains in human mucin peptides mapped to 11p15.

A;Reference number: S35047; MUID:93343858; PMID:7916618

A;Accession: S35048

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-330 <DUF>

Query Match 53.1%; Score 43; DB 2; Length 330;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 PPPPQETVT 15
: ||||| :|
Db 186 PPPPQEQLT 194

RESULT 11

JC5505

DNA helicase Q1 interacting protein 1 - human

N;Alternate names: nuclear localization signal receptor importin-alpha homolog

C;Species: Homo sapiens (man)

C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 20-Jun-2000

C;Accession: JC5505

R;Seki, T.; Tada, S.; Katada, T.; Enomoto, T.

Biochem. Biophys. Res. Commun. 234, 48-53, 1997

A;Title: Cloning of a cDNA encoding a novel importin-alpha homologue, Qip1: Discrimination of Qip1 and Rchl from hSrp1 by their ability to interact with DNA helicase Q1/RecQL.
 A;Reference number: JC5505; MUID:97312517; PMID:9168958
 A;Accession: JC5505
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-521 <SEK>
 A;Cross-references: DDBJ:AB002533; NID:g1944124; PIDN:BAA19546.1; PID:g1944125
 C;Comment: This protein interacts with the nuclear localizaiton signal motif of SV40 T antigen and of DNA helicase Q1. It plays a role in the process of nuclear protein transport into the nucleus.
 C;Superfamily: pendulin
 F;11-51/Region: importin-beta binding #status predicted

Query Match 53.1%; Score 43; DB 2; Length 521;
 Best Local Similarity 70.0%; Pred. No. 60;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 KFPPPPQETV 14
 | |||| ||:
 Db 231 KDPPPPMETI 240

RESULT 12

S35047

mucin JUL7 - human

C;Species: Homo sapiens (man)

C;Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 08-Oct-1999

C;Accession: S35047

R;Dufosse, J.; Porchet, N.; Audie, J.P.; Guyonnet Duperat, V.; Laine, A.; van-Seuningen, I.; Marrakchi, S.; Degand, P.; Aubert, J.P.
 Biochem. J. 293, 329-337, 1993

A;Title: Degenerate 87-base-pair tandem repeats create hydrophilic/hydrophobic alternating domains in human mucin peptides mapped to 11p15.

A;Reference number: S35047; MUID:93343858; PMID:7916618

A;Accession: S35047

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-543 <DUF>

A;Cross-references: EMBL:X74370; NID:g407081; PIDN:CAA52408.1; PID:g407082

A;Note: the authors translated the codon AAA for residue 63 as Gln and CGG for residue 379 as Pro

A;Note: the amino acid sequence from Fig. 2 is inconsistent with the nucleotide sequence from Fig. 1 in having the additional residue Thr after 329-Thr

Query Match 53.1%; Score 43; DB 2; Length 543;
 Best Local Similarity 77.8%; Pred. No. 63;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 PPPPQETVT 15
 ||||| :|
 Db 361 PPPPQEQLT 369

RESULT 13

T39990

hypothetical protein SPBC25D12.02c - fission yeast (*Schizosaccharomyces pombe*)
 C;Species: *Schizosaccharomyces pombe*
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C;Accession: T39990
 R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Whithead, S.; Chillingworth, T.; Churcher, C.M.
 submitted to the EMBL Data Library, July 1998
 A;Reference number: Z21897
 A;Accession: T39990
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-599 <LYN>
 A;Cross-references: EMBL:AL031158; PIDN:CAA20098.1; GSPDB:GN00067;
 SPDB:SPBC25D12.02c
 A;Experimental source: strain 972h-; cosmid c25D12
 C;Genetics:
 A;Gene: SPDB:SPBC25D12.02c
 A;Map position: 2
 A;Introns: 48/2

Query Match 53.1%; Score 43; DB 2; Length 599;
 Best Local Similarity 72.7%; Pred. No. 71;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 KFPPPPQETVT 15
 ||||| | ||
 Db 415 KFPPPSQTLVT 425

RESULT 14

S71110
 abcl protein homolog precursor - fission yeast (*Schizosaccharomyces pombe*)
 C;Species: *Schizosaccharomyces pombe*
 C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 31-Jan-2000
 C;Accession: S71111; T40121; S71110
 R;Bonnefoy, N.; Kermorgant, M.; Brivet-Chevillotte, P.; Dujardin, G.
 Mol. Gen. Genet. 251, 204-210, 1996
 A;Title: Cloning by functional complementation, and inactivation, of the
Schizosaccharomyces pombe homologue of the *Saccharomyces cerevisiae* gene ABC1.
 A;Reference number: S71111; MUID:96242153; PMID:8668131
 A;Accession: S71111
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 177-610 <BOW>
 A;Cross-references: EMBL:X91616; NID:g1514640; PIDN:CAA62818.1; PID:e199846;
 PID:g1514641
 R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Taylor, K.; Harris, D.
 submitted to the EMBL Data Library, September 1998
 A;Reference number: Z21906
 A;Accession: T40121
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-610 <WOO>
 A;Cross-references: EMBL:AL031788; PIDN:CAA21176.1; GSPDB:GN00067;
 SPDB:SPBC2D10.18
 A;Experimental source: strain 972h-; cosmid c2D10
 C;Genetics:

A;Gene: abcl; SPBC2D10.18

A;Map position: 2

Query Match 53.1%; Score 43; DB 2; Length 610;
Best Local Similarity 46.2%; Pred. No. 72;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
:|::: |||:|
Db 562 MLDLRLQPPPEET 574

RESULT 15

T09941

transcription factor VP1 - Craterostigma plantagineum

N;Alternate names: VP1 homolog

C;Species: Craterostigma plantagineum

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 18-Aug-2000

C;Accession: T09941

R;Chandler, J.W.; Bartels, D.

Mol. Gen. Genet. 256, 539-546, 1997

A;Title: Structure and function of the vp1 gene homologue from the resurrection plant Craterostigma plantagineum Hochst.

A;Reference number: Z16899; MUID:98075359; PMID:9413438

A;Accession: T09941

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-688 <CHA>

A;Cross-references: EMBL:AJ000552; NID:g2288898; PIDN:CAA04184.1; PID:g2288899

A;Experimental source: seed

C;Genetics:

A;Gene: vp1

C;Function:

A;Description: seed-specific transcriptional activator

C;Superfamily: rice transcription factor VP1

C;Keywords: DNA binding; seed; transcription factor; transcription regulation

Query Match 53.1%; Score 43; DB 2; Length 688;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 PPPPQET 13
|||||||
Db 321 PPPPQET 327

Search completed: August 24, 2004, 15:52:56

Job time : 16.5522 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:51:19 ; Search time 54.291 Seconds
(without alignments)
86.825 Million cell updates/sec

Title: US-09-641-801-7
Perfect score: 81
Sequence: 1 VLEMKFPPPPQETVT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 314255058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query					
No.	Score	Match	Length	DB	ID	Description

1	81	100.0	15	14	US-10-281-652-7	Sequence 7, Appli
2	51	63.0	289	12	US-10-412-699B-314	Sequence 314, App
3	51	63.0	289	15	US-10-225-068-104	Sequence 104, App
4	51	63.0	289	15	US-10-374-780A-2014	Sequence 2014, Ap
5	51	63.0	319	10	US-09-533-029-38	Sequence 38, Appl
6	51	63.0	319	14	US-10-286-264-22	Sequence 22, Appl
7	51	63.0	319	14	US-10-278-536-148	Sequence 148, App
8	46	56.8	1196	15	US-10-369-493-3649	Sequence 3649, Ap
9	45	55.6	56	12	US-10-424-599-158309	Sequence 158309,
10	45	55.6	78	9	US-09-864-761-37299	Sequence 37299, A
11	44	54.3	61	12	US-10-424-599-220767	Sequence 220767,
12	44	54.3	95	16	US-10-437-963-127360	Sequence 127360,
13	44	54.3	102	12	US-10-424-599-170529	Sequence 170529,
14	44	54.3	111	16	US-10-437-963-176514	Sequence 176514,
15	44	54.3	137	12	US-10-424-599-265006	Sequence 265006,
16	44	54.3	158	12	US-10-424-599-161423	Sequence 161423,
17	44	54.3	166	12	US-10-225-066A-86	Sequence 86, Appl
18	44	54.3	166	15	US-10-225-067-10	Sequence 10, Appl
19	44	54.3	166	15	US-10-374-780A-2810	Sequence 2810, Ap
20	44	54.3	174	12	US-10-425-114-40629	Sequence 40629, A
21	44	54.3	289	14	US-10-011-585A-167	Sequence 167, App
22	44	54.3	361	9	US-09-864-761-33614	Sequence 33614, A
23	44	54.3	361	9	US-09-864-761-34141	Sequence 34141, A
24	44	54.3	420	16	US-10-437-963-106810	Sequence 106810,
25	44	54.3	515	16	US-10-437-963-169499	Sequence 169499,
26	44	54.3	521	12	US-10-231-956A-154	Sequence 154, App
27	44	54.3	529	9	US-09-925-301-866	Sequence 866, App
28	44	54.3	634	16	US-10-437-963-196028	Sequence 196028,
29	44	54.3	1047	14	US-10-029-386-33522	Sequence 33522, A
30	44	54.3	1087	14	US-10-029-386-32407	Sequence 32407, A
31	43	53.1	157	12	US-10-424-599-226767	Sequence 226767,
32	43	53.1	158	12	US-10-425-114-37363	Sequence 37363, A
33	43	53.1	176	16	US-10-437-963-115326	Sequence 115326,
34	43	53.1	201	16	US-10-437-963-103502	Sequence 103502,
35	43	53.1	214	12	US-10-424-599-240944	Sequence 240944,
36	43	53.1	219	9	US-09-925-301-1039	Sequence 1039, Ap
37	43	53.1	219	11	US-09-997-003-35	Sequence 35, Appl
38	43	53.1	219	11	US-09-997-003-47	Sequence 47, Appl
39	43	53.1	329	14	US-10-106-698-4702	Sequence 4702, Ap
40	43	53.1	337	16	US-10-437-963-181432	Sequence 181432,
41	43	53.1	414	12	US-10-424-599-213840	Sequence 213840,
42	43	53.1	417	16	US-10-437-963-143835	Sequence 143835,
43	43	53.1	438	16	US-10-437-963-146562	Sequence 146562,
44	43	53.1	462	12	US-10-425-114-64426	Sequence 64426, A
45	43	53.1	472	12	US-10-425-114-41069	Sequence 41069, A

ALIGNMENTS

RESULT 1

US-10-281-652-7

; Sequence 7, Application US/10281652

; Publication No. US20030091606A1

; GENERAL INFORMATION:

; APPLICANT: STANTON, G. John

```

; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDOGH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND
; TITLE OF INVENTION: ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/10/281,652
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/09/641,803
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,310
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-10-281-652-7

```

```

Query Match          100.0%; Score 81; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 VLEMKFPPPPQETVT 15
        |||
Db      1 VLEMKFPPPPQETVT 15

```

RESULT 2

US-10-412-699B-314

```

; Sequence 314, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP

```

```

; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 314
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502
US-10-412-699B-314

```

```

Query Match          63.0%; Score 51; DB 12; Length 289;
Best Local Similarity 69.2%; Pred. No. 38;
Matches      9; Conservative 1; Mismatches      3; Indels      0; Gaps      0;

```

```

Qy      1 VLEMKFPPPPQET 13
        | | | | | | | |
Db      184 VTEMVMPPPPQQT 196

```

```

RESULT 3
US-10-225-068-104
; Sequence 104, Application US/10225068
; Publication No. US20030217383A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Adam, Luc J.
; APPLICANT: Dubell, Arnold T.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang

```

```

; APPLICANT: Broun, Pierre E.
; TITLE OF INVENTION: STRESS-RELATED POLYNUCLEOTIDES AND
; TITLE OF INVENTION: POLYPEPTIDES IN PLANTS
; FILE REFERENCE: 514442002040
; CURRENT APPLICATION NUMBER: US/10/225,068
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (10)...(155)
; OTHER INFORMATION: Conserved domain
US-10-225-068-104

```

```

Query Match          63.0%; Score 51; DB 15; Length 289;
Best Local Similarity 69.2%; Pred. No. 38;
Matches      9; Conservative 1; Mismatches      3; Indels      0; Gaps      0;

```

```

Qy      1 VLEMKFPPPPQET 13
        | ||  |||||:|
Db      184 VTEMVMPPPPQQT 196

```

RESULT 4

US-10-374-780A-2014

```

; Sequence 2014, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP

```

```

; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2014
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502 Paralogous to G519
US-10-374-780A-2014

```

```

Query Match          63.0%; Score 51; DB 15; Length 289;
Best Local Similarity 69.2%; Pred. No. 38;
Matches      9; Conservative 1; Mismatches      3; Indels      0; Gaps      0;

```

```

QY      1 VLEMKFPPPPQET 13
        | | | | | | | |
Db      184 VTEMVMPPPPQQT 196

```

```

RESULT 5
US-09-533-029-38
; Sequence 38, Application US/09533029
; Publication No. US20030046723A1
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne

```

```

; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502
US-09-533-029-38

```

```

Query Match          63.0%; Score 51; DB 10; Length 319;
Best Local Similarity 69.2%; Pred. No. 42;
Matches      9; Conservative 1; Mismatches      3; Indels      0; Gaps      0;

```

```

Qy      1 VLEMKFPPPPQET 13
        | | | | | | | |
Db      184 VTEMVMPPPPQQT 196

```

RESULT 6

```

US-10-286-264-22
; Sequence 22, Application US/10286264
; Publication No. US20030093837A1
; GENERAL INFORMATION:
; APPLICANT: Keddie, James
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Zhang, James
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Pineda, Omaira
; APPLICANT: Heard, Jacqueline
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Adam, Luc
; APPLICANT: Broun, Pierre
; APPLICANT: Reuber, Lynne
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Samaha, Raymond
; TITLE OF INVENTION: POLYNUCLEOTIDES FOR SEED TRAIT ALTERATION
; FILE REFERENCE: MBI-008
; CURRENT APPLICATION NUMBER: US/10/286,264
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502

```

US-10-286-264-22

Query Match 63.0%; Score 51; DB 14; Length 319;
Best Local Similarity 69.2%; Pred. No. 42;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
| || |||||:|
Db 184 VTEMVMPPPPQQT 196

RESULT 7

US-10-278-536-148
; Sequence 148, Application US/10278536
; Publication No. US20030131386A1
; GENERAL INFORMATION:
; APPLICANT: Samaha, Raymond
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Keddie, James
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Adam, Luc
; APPLICANT: Broun, Pierre
; TITLE OF INVENTION: STRESS-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-011
; CURRENT APPLICATION NUMBER: US/10/278,536
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 148
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502
US-10-278-536-148

Query Match 63.0%; Score 51; DB 14; Length 319;
Best Local Similarity 69.2%; Pred. No. 42;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
| || |||||:|
Db 184 VTEMVMPPPPQQT 196

RESULT 8

US-10-369-493-3649
; Sequence 3649, Application US/10369493
; Publication No. US20030233675A1


```

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION
OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3649
; LENGTH: 1196
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1196)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3649

```

```

Query Match          56.8%; Score 46; DB 15; Length 1196;
Best Local Similarity 57.1%; Pred. No. 6.8e+02;
Matches      8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy      2 LEMKFPPPPQETVT 15
        | : ||||| | |
Db      1089 LSLSFPPPPPTPT 1102

```

RESULT 9

```

US-10-424-599-158309
; Sequence 158309, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 158309
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_113972C.1.pep
US-10-424-599-158309

```

Query Match 55.6%; Score 45; DB 12; Length 56;
Best Local Similarity 58.3%; Pred. No. 51;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQE 12
||::| |||:::
Db 16 VLKIKIPPPPRK 27

RESULT 10

US-09-864-761-37299

; Sequence 37299, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

```

; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37299
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010906.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
US-09-864-761-37299

```

```

Query Match          55.6%; Score 45; DB 9; Length 78;
Best Local Similarity 70.0%; Pred. No. 69;
Matches      7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      4 MKFP PPPQET 13
        :| |||||:|
Db      39 LKTP PPPQQT 48

```

RESULT 11

```

US-10-424-599-220767
; Sequence 220767, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 220767
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_41383C.1.pep
US-10-424-599-220767

```

Query Match 54.3%; Score 44; DB 12; Length 61;
Best Local Similarity 87.5%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 FPPPPQET 13
||||| ||
Db 14 FPPPPSET 21

RESULT 12

US-10-437-963-127360
; Sequence 127360, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 127360
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2981C.1.pep
US-10-437-963-127360

Query Match 54.3%; Score 44; DB 16; Length 95;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EMKFPPPPQE 12
| | |||||
Db 5 ERKKPPPPQE 14

RESULT 13

US-10-424-599-170529
; Sequence 170529, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 170529
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_125002C.1.pep
US-10-424-599-170529

Query Match 54.3%; Score 44; DB 12; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 FPPPPQE 12
| | | | |
Db 51 FPPPPQE 57

RESULT 14

US-10-437-963-176514
; Sequence 176514, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 176514
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_74254C.1.pep
US-10-437-963-176514

Query Match 54.3%; Score 44; DB 16; Length 111;
Best Local Similarity 72.7%; Pred. No. 1.3e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 KFPPPPQETVT 15
| | | | : | |
Db 21 KSPSPQQTVT 31

RESULT 15

US-10-424-599-265006

; Sequence 265006, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 265006

; LENGTH: 137

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(137)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_8131C.1.pep

US-10-424-599-265006

Query Match 54.3%; Score 44; DB 12; Length 137;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 FPPPPQE 12

|||||||

Db 112 FPPPPQE 118

Search completed: August 24, 2004, 16:41:25

Job time : 57.291 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:23:00 ; Search time 46.3433 Seconds
(without alignments)
102.124 Million cell updates/sec

Title: US-09-641-801-7
Perfect score: 81
Sequence: 1 VLEMKFPPPPQETVT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match Length DB	ID			Description

1	51	63.0	229	10	Q39013	Q39013 arabidopsis
2	51	63.0	289	10	Q9LQ85	Q9lq85 arabidopsis
3	49	60.5	180	10	O82297	O82297 arabidopsis
4	47	58.0	375	11	Q99J35	Q99j35 mus musculu
5	47	58.0	526	16	Q92KS7	Q92ks7 rhizobium m
6	46	56.8	148	16	Q9I5B7	Q9i5b7 pseudomonas
7	46	56.8	272	2	Q02304	Q02304 pseudomonas
8	46	56.8	4587	11	Q9QXA3	Q9qxa3 mus musculu
9	45	55.6	114	5	Q9W0Z6	Q9w0z6 drosophila
10	45	55.6	389	2	Q9F183	Q9f183 alcaligenes
11	44	54.3	166	10	Q9C5I3	Q9c5i3 arabidopsis
12	44	54.3	172	4	Q8IYL7	Q8iyl7 homo sapien
13	44	54.3	201	6	Q9GK20	Q9gk20 sus scrofa
14	44	54.3	204	10	Q9SGN4	Q9sgn4 arabidopsis
15	44	54.3	288	4	Q9NWJ6	Q9nwj6 homo sapien
16	44	54.3	337	11	Q9CT07	Q9ct07 mus musculu
17	44	54.3	520	13	Q8JFT7	Q8jft7 brachydanio
18	44	54.3	520	13	Q7SXT4	Q7sxt4 brachydanio
19	44	54.3	521	4	Q8IYQ9	Q8iyq9 homo sapien
20	44	54.3	528	10	Q9ZPL6	Q9zpl6 nicotiana t
21	44	54.3	642	12	P87502	P87502 human adeno
22	44	54.3	665	4	Q8IWS8	Q8iws8 homo sapien
23	44	54.3	666	4	Q8IZZ7	Q8izz7 homo sapien
24	44	54.3	666	4	Q7Z5R6	Q7z5r6 homo sapien
25	44	54.3	1033	5	Q8T2L2	Q8t2l2 dictyosteli
26	43	53.1	61	12	Q91FX0	Q91fx0 chilo iride
27	43	53.1	160	10	Q93W21	Q93w21 oryza sativ
28	43	53.1	166	6	Q9GK22	Q9gk22 sus scrofa
29	43	53.1	171	11	Q63401	Q63401 rattus norv
30	43	53.1	190	10	Q8GSB4	Q8gsb4 oryza sativ
31	43	53.1	196	5	Q23610	Q23610 caenorhabdi
32	43	53.1	330	4	Q14882	Q14882 homo sapien
33	43	53.1	380	10	Q7X9Z9	Q7x9z9 fagopyrum c
34	43	53.1	492	16	Q93HH8	Q93hh8 streptomyce
35	43	53.1	513	10	Q9FGF3	Q9fgf3 arabidopsis
36	43	53.1	521	13	Q7ZVY9	Q7zvy9 brachydanio
37	43	53.1	543	4	Q14879	Q14879 homo sapien
38	43	53.1	599	3	O74354	O74354 schizosacch
39	43	53.1	628	10	Q9SF51	Q9sf51 arabidopsis
40	43	53.1	640	5	Q95U09	Q95u09 drosophila
41	43	53.1	688	10	O23774	O23774 craterostig
42	43	53.1	689	11	Q80VA2	Q80va2 mus musculu
43	43	53.1	709	5	Q9XWN6	Q9xwn6 caenorhabdi
44	43	53.1	802	16	Q7UWT0	Q7uwt0 rhodopirell
45	43	53.1	810	5	Q9VJ08	Q9vj08 drosophila

ALIGNMENTS

RESULT 1

Q39013

ID Q39013 PRELIMINARY; PRT; 229 AA.

AC Q39013;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE ATAF1 protein.
 GN ATAF1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rueth J., Schweyen R., Hirt H.;
 RL Submitted (AUG-1993) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; X74755; CAA52771.1; -.
 DR PIR; S37101; S37101.
 DR InterPro; IPR003441; NAM.
 DR Pfam; PF02365; NAM; 1.
 SQ SEQUENCE 229 AA; 26031 MW; 474CE31D84E68467 CRC64;

Query Match 63.0%; Score 51; DB 10; Length 229;
 Best Local Similarity 69.2%; Pred. No. 1.2;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
 | | | | | | | | | |
 Db 124 VTEMVMPPPPQQT 136

RESULT 2

Q9LQ85

ID Q9LQ85 PRELIMINARY; PRT; 289 AA.
 AC Q9LQ85;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE T1N6.12 protein (OsNAC6 protein-like protein).
 GN T1N6.12.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Liu S., Chan A., Yu G., Lee J., Lenz C., Pham P., Sakano H.,
 RA Toriumi M., Chin C., Chiou J., Choi E., Chung M., Gonzalez A.,
 RA Howng B., Liu A., Vaysberg M., Altafi H., Brooks S., Buehler E.,
 RA Chao Q., Conn L., Conway A., Hansen N., Johnson-Hopson C., Khan S.,
 RA Kim C., Lam B., Miranda M., Nguyen M., Palm C., Shinn P.,
 RA Southwick A., Davis R., Ecker J., Federspiel N., Theologis A.;
 RT "The sequence of BAC T1N6 from Arabidopsis thaliana chromosome 1."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Chen H., Cheuk R.,
 RA Chung M.K., Kim C., Lin J., Liu S.X., Pham P.K., Sakano H., Shinn P.,
 RA Yamada K., Ecker J., Theologis A., Davis R.W.;

RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AC009273; AAF78403.1; -.
 DR EMBL; AF370617; AAK43936.1; -.
 DR PIR; E86148; E86148.
 DR InterPro; IPR003441; NAM.
 DR Pfam; PF02365; NAM; 1.
 SQ SEQUENCE 289 AA; 32922 MW; DDAC62A101B6CE4B CRC64;

Query Match 63.0%; Score 51; DB 10; Length 289;
 Best Local Similarity 69.2%; Pred. No. 1.5;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
 | || |||||:|
 Db 184 VTEMVMPPPPQQT 196

RESULT 3

O82297

ID O82297 PRELIMINARY; PRT; 180 AA.
 AC O82297;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE At2g35430 protein.
 GN AT2G35430.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana."
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AC005314; AAC36178.1; -.
 DR PIR; E84768; E84768.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR000571; Znf_CCCH.
 DR Pfam; PF00642; zf-CCCH; 2.
 DR SMART; SM00356; Znf_C3H1; 2.
 SQ SEQUENCE 180 AA; 19991 MW; 9039E7F8120E4FF2 CRC64;

Query Match 60.5%; Score 49; DB 10; Length 180;
Best Local Similarity 58.8%; Pred. No. 2.1;
Matches 10; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

Qy 3 EMKFPPPP---QETVT 15
|:: |||| ||||
Db 100 ELRLPPPPPPNWQETVT 116

RESULT 4

Q99J35

ID Q99J35 PRELIMINARY; PRT; 375 AA.
AC Q99J35;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN 2410043F08RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: CONTAINS 1 LIM DOMAIN. THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
DR EMBL; BC004777; AAH04777.1; -.
DR MGD; MGI:1921452; 2410043F08Rik.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 3.
DR ProDom; PD000094; LIM; 3.
DR SMART; SM00132; LIM; 3.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR PROSITE; PS00478; LIM_DOMAIN_1; 3.
DR PROSITE; PS50023; LIM_DOMAIN_2; 1.
KW Hypothetical protein; LIM domain; Metal-binding; Zinc.
SQ SEQUENCE 375 AA; 41026 MW; 601DF99A8EB1BB80 CRC64;

Query Match 58.0%; Score 47; DB 11; Length 375;
Best Local Similarity 61.5%; Pred. No. 9.2;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 EMKFPPPPQETVT 15
| : ||||:| ||
Db 162 EEELPPPPEEPVT 174

RESULT 5

Q92KS7

ID Q92KS7 PRELIMINARY; PRT; 526 AA.

AC Q92KS7;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein R00256.
 GN R00256 OR SMC00336.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL; AL591783; CAC41693.1; -.
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
 DR InterPro; IPR003454; MmoB_DmpM.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 526 AA; 55373 MW; 6D657535288CB651 CRC64;

Query Match 58.0%; Score 47; DB 16; Length 526;
 Best Local Similarity 75.0%; Pred. No. 13;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEMKFP PPPQET 13
 | :|||| ||||
 Db 68 LILKFPPLPQET 79

RESULT 6

Q9I5B7

ID Q9I5B7 PRELIMINARY; PRT; 148 AA.
 AC Q9I5B7;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein PA0822.
 GN PA0822.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL; AE004517; AAG04211.1; -.
 DR PIR; F83542; F83542.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 148 AA; 16617 MW; EB8056036D1CBA4F CRC64;

Query Match 56.8%; Score 46; DB 16; Length 148;
 Best Local Similarity 80.0%; Pred. No. 5.4;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPP 10
 |:||||| |
 Db 69 VVEMKFPPDP 78

RESULT 7

Q02304

ID Q02304 PRELIMINARY; PRT; 272 AA.
 AC Q02304;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE ORF3 protein.
 OS *Pseudomonas* sp.
 OC Bacteria; Proteobacteria.
 OX NCBI_TaxID=306;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92339462; PubMed=1633822;
 RA Altenschmidt U., Bokranz M., Fuchs G.;
 RT "Novel aerobic 2-aminobenzoate metabolism. Nucleotide sequence of the
 RT plasmid carrying the gene for the flavoprotein 2-aminobenzoyl-CoA
 RT monooxygenase/reductase in a denitrifying *Pseudomonas* sp."
 RL Eur. J. Biochem. 207:715-722(1992).
 DR EMBL; X66604; CAA47169.1; -.
 DR PIR; S24375; S24375.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0005840; C:ribosome; IEA.
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR001209; Ribosomal_S14.
 DR InterPro; IPR008900; Zot.
 DR Pfam; PF05707; Zot; 1.
 DR PROSITE; PS00527; RIBOSOMAL_S14; 1.
 SQ SEQUENCE 272 AA; 30631 MW; 1776F7C19A429050 CRC64;

Query Match 56.8%; Score 46; DB 2; Length 272;
 Best Local Similarity 58.3%; Pred. No. 9.9;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LEMKFPPPPQET 13

Db | :|||| :|
153 LRRRFPPPPDDT 164

RESULT 8

Q9QXA3

ID Q9QXA3 PRELIMINARY; PRT; 4587 AA.
AC Q9QXA3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mouse fat 1 cadherin (Fragment).
GN FATH OR MFAT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=20203461; PubMed=10741417;
RA Cox B.T.M., Hadjantonakis A.K., Collins J., Magee A.I.;
RT "Cloning and expression throughout mouse development of mfat1 a
RT homologue of the Drosophila tumour suppressor gene fat.";
RL Dev. Dyn. 217:233-240(2000).
DR EMBL; AJ250768; CAB65271.1; -.
DR HSSP; P08709; 1BF9.
DR MGD; MGI:109168; Fath.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005194; F:cell adhesion molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR001791; Laminin_G.
DR Pfam; PF00028; cadherin; 33.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF00054; laminin_G; 1.
DR PRINTS; PR00205; CADHERIN.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00112; CA; 31.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00282; LamG; 1.
DR PROSITE; PS00232; CADHERIN_1; 17.
DR PROSITE; PS50268; CADHERIN_2; 33.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS50025; LAM_G_DOMAIN; 1.
KW Calcium; Calcium-binding; Cell adhesion; Glycoprotein.
FT NON TER 4587 4587
SQ SEQUENCE 4587 AA; 506036 MW; 4D3F23BB05127CB4 CRC64;

Query Match 56.8%; Score 46; DB 11; Length 4587;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEMKFPPPPQE 12
:| ||||:|
Db 4430 IESDFPPPPPEE 4440

RESULT 9

Q9W0Z6

ID Q9W0Z6 PRELIMINARY; PRT; 114 AA.
AC Q9W0Z6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG15859 protein.
GN CG15859.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003465; AAF47282.1; -.
 DR FlyBase; FBgn0035071; CG15859.
 SQ SEQUENCE 114 AA; 11621 MW; C9481842A9EDE0BB CRC64;

Query Match 55.6%; Score 45; DB 5; Length 114;
 Best Local Similarity 77.8%; Pred. No. 6.1;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 7 PPPPQETVT 15
 |||||:|:
 Db 101 PPPPQDTV 109

RESULT 10

Q9F183

ID Q9F183 PRELIMINARY; PRT; 389 AA.
 AC Q9F183;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS *Alcaligenes eutrophus* (*Ralstonia eutropha*).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; *Ralstonia*.
 OX NCBI_TaxID=510;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=335T;
 RA Hinner I.-S., Buerger S., Schloemann M.;
 RT "Characterization of a second gene cluster coding for enzymes of
 RT catechol catabolism in *Ralstonia eutropha* 335T.";
 RL Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF042280; AAG42031.1; -.
 DR InterPro; IPR001466; Beta_lactamase.
 DR Pfam; PF00144; beta-lactamase; 1.
 KW Hypothetical protein.
 FT NON_TER 389 389
 SQ SEQUENCE 389 AA; 42244 MW; 600D57AE93338113 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 389;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 LEMKFPPPPQETVT 15
 |: ||||| : :|
 Db 24 LDAGFPPPPDKLIT 37

RESULT 11

Q9C5I3

ID Q9C5I3 PRELIMINARY; PRT; 166 AA.
AC Q9C5I3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative ethylene responsive element binding factor 4 protein (ERF
DE domain protein11) (Ethylene-responsive element binding factor,
DE putative).
GN F3M18.20 OR ATERF11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full Length cDNA of gene F3M18.20 (GI:6560756).";
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21380424; PubMed=11487705;
RA Ohta M., Matsui K., Hiratsu K., Shinshi H., Ohme-Takagi M.;
RT "Repression Domains of Class II ERF Transcriptional Repressors Share
RT an Essential Motif for Active Repression.";
RL Plant Cell 13:1959-1968(2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Dale J.M., Gibson H.A., Goldsmith A.D., Jiang P.X., Lee J.M.,
RA Quach H.L., Tang C., Toriumi M., Yu G., Bowser L., Carninci P.,
RA Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full Length cDNA of gene F3M18.20 (GI:6560756).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 0:0-0(2002).
RN [5]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;

RT "Full-Length cDNA from Arabidopsis thaliana.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF360232; AAK25942.1; -.
 DR EMBL; AB055882; BAB62911.1; -.
 DR EMBL; AY040032; AAK64090.1; -.
 DR EMBL; AY086783; AAM63833.1; -.
 DR HSSP; O80337; 2GCC.
 DR TRANSFAC; T05479; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR PRINTS; PR00367; ETHRSPELEMNT.
 DR ProDom; PD001423; TF_ERF; 1.
 SQ SEQUENCE 166 AA; 18612 MW; A327522DB20E5984 CRC64;

Query Match 54.3%; Score 44; DB 10; Length 166;
 Best Local Similarity 60.0%; Pred. No. 13;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LEMKFPPPPQ 11
 |:: ||||:
 Db 156 LDLNFPPPPPE 165

RESULT 12

Q8IYL7

ID Q8IYL7 PRELIMINARY; PRT; 172 AA.
 AC Q8IYL7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to proline-rich protein 48.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; BC035636; AAH35636.1; -.
 SQ SEQUENCE 172 AA; 18685 MW; 21314883C4E8C7B9 CRC64;

Query Match 54.3%; Score 44; DB 4; Length 172;
 Best Local Similarity 46.7%; Pred. No. 14;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
 ||:: |||| |::
 Db 136 VLDLPLPPPPPEPLS 150

RESULT 13

Q9GK20

ID Q9GK20 PRELIMINARY; PRT; 201 AA.
AC Q9GK20;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Karyopherin alpha 3 (Fragment).
GN KPNA3.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA VanVickle S.J., Cabot R.A., Prather R.S.;
RT "Porcine karyopherin (importin) alpha 3 (KPNA3).";
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY010069; AAG42107.2; -.
DR HSSP; Q02821; 1BK5.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 5.
DR SMART; SM00185; ARM; 5.
DR PROSITE; PS50176; ARM_REPEAT; 2.
FT NON_TER 1 1
FT NON_TER 201 201
SQ SEQUENCE 201 AA; 22259 MW; FAE90BEB5D6489D1 CRC64;

Query Match 54.3%; Score 44; DB 6; Length 201;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 KFPPPPQETV 14
| |||| |||
Db 78 KDPPPPMETV 87

RESULT 14

Q9SGN4

ID Q9SGN4 PRELIMINARY; PRT; 204 AA.
AC Q9SGN4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE F3M18.20.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,

RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F3M18 from chromosome
 RT I.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AC010155; AAF16756.1; -.
 DR PIR; B86410; B86410.
 DR HSSP; O80337; 2GCC.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR PRINTS; PR00367; ETHRSPELEMNT.
 DR ProDom; PD001423; TF_ERF; 1.
 DR SMART; SM00380; AP2; 1.
 SQ SEQUENCE 204 AA; 23213 MW; C889B6C0D633BB16 CRC64;

Query Match 54.3%; Score 44; DB 10; Length 204;
 Best Local Similarity 60.0%; Pred. No. 16;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LEMKFPPPPQ 11
 |:: |||||:
 Db 194 LDLNFPPPPPE 203

RESULT 15

Q9NWJ6

ID Q9NWJ6 PRELIMINARY; PRT; 288 AA.
 AC Q9NWJ6;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein FLJ20805.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Adipose tissue;
 RA Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,
 RA Hirao M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,
 RA Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AK000812; BAA91384.1; -.
 DR GO; GO:0007218; P:neuropeptide signaling pathway; IEA.
 DR InterPro; IPR000159; RA_domain.
 DR Pfam; PF00788; RA; 1.
 DR SMART; SM00314; RA; 1.
 DR PROSITE; PS50200; RA; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 288 AA; 32179 MW; 6EBD2A490134FD43 CRC64;

Query Match 54.3%; Score 44; DB 4; Length 288;
 Best Local Similarity 46.7%; Pred. No. 23;

Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
||:: |||| | ::
Db 136 VLDLPLPPPPPEPLS 150

Search completed: August 24, 2004, 15:50:51
Job time : 49.3433 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 14:57:04 ; Search time 8.0597 Seconds
(without alignments)
96.908 Million cell updates/sec

Title: US-09-641-801-7
Perfect score: 81
Sequence: 1 VLEMKFPPPPQETVT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	44	54.3	521	1	IMA3_HUMAN	O00505 homo sapien
2	44	54.3	521	1	IMA3_MOUSE	O35344 mus musculu
3	44	54.3	2426	1	SON_HUMAN	P18583 homo sapien
4	43	53.1	521	1	IMA4_HUMAN	O00629 homo sapien
5	43	53.1	521	1	IMA4_MOUSE	O35343 mus musculu
6	43	53.1	610	1	ABCI_SCHPO	Q92338 schizosacch
7	43	53.1	4590	1	FATH_HUMAN	Q14517 homo sapien
8	42.5	52.5	690	1	CRN1_MOUSE	Q9cqc1 mus musculu
9	42.5	52.5	848	1	CRN1_HUMAN	Q9bzbj0 homo sapien
10	42	51.9	182	1	FTN_DROME	P35554 drosophila
11	42	51.9	291	1	CD20_MOUSE	P19437 mus musculu
12	42	51.9	404	1	RHB4_HUMAN	P58872 homo sapien
13	42	51.9	475	1	LEGU_CANEN	P49046 canavalia e
14	41.5	51.2	460	1	MOC3_HUMAN	O95396 homo sapien
15	41	50.6	208	1	YK69_CAEEL	P34347 caenorhabdi
16	41	50.6	213	1	HS27_DROME	P02518 drosophila
17	41	50.6	373	1	FBLP_HUMAN	Q8wup2 homo sapien

18	41	50.6	493	1	VPE2_PHAVU	O24326	phaseolus v
19	41	50.6	1224	1	ABL1_CAEEL	P03949	caenorhabdi
20	41	50.6	1250	1	SSD1_YEAST	P24276	saccharomyc
21	41	50.6	2845	1	APC_MOUSE	Q61315	mus musculu
22	40.5	50.0	1051	1	PTK7_CHICK	Q91048	gallus gall
23	40.5	50.0	1070	1	PTK7_HUMAN	Q13308	homo sapien
24	40	49.4	256	1	RK21_SPIOL	P24613	spinacia ol
25	40	49.4	390	1	RBM9_HUMAN	O43251	homo sapien
26	40	49.4	550	1	VGLE_HSVEB	P18552	equine herp
27	40	49.4	552	1	VGLE_HSVEL	P24380	equine herp
28	40	49.4	618	1	PDA4_CAEEL	P34329	caenorhabdi
29	40	49.4	800	1	P5CS_CAEEL	P54889	c probable
30	40	49.4	863	1	HCN2_MOUSE	O88703	mus musculu
31	40	49.4	1214	1	BRF1_HUMAN	P55201	homo sapien
32	40	49.4	2150	1	POLG_HRV2	P04936	h genome po
33	39	48.1	88	1	Y419_NEIMA	Q9jqp5	neisseria m
34	39	48.1	322	1	UL15_HCMVA	P16844	human cytom
35	39	48.1	466	1	NRAM_INBLE	P03474	influenza b
36	39	48.1	536	1	NHR7_CAEEL	Q20765	caenorhabdi
37	39	48.1	592	1	ABP1_YEAST	P15891	saccharomyc
38	39	48.1	598	1	YDW3_SCHPO	O13910	schizosacch
39	39	48.1	639	1	ACTA_LISMO	P33379	listeria mo
40	39	48.1	662	1	ERF2_SCHPO	O74718	schizosacch
41	39	48.1	730	1	ABB2_HUMAN	Q92870	homo sapien
42	38.5	47.5	465	1	RSP4_CHLRE	Q01656	chlamydomon
43	38	46.9	208	1	SPC3_STRPU	P16537	strongyloce
44	38	46.9	220	1	NO30_PHAVU	Q41121	phaseolus v
45	38	46.9	396	1	YNE2_YEAST	P53958	saccharomyc

ALIGNMENTS

RESULT 1

IMA3_HUMAN

ID IMA3_HUMAN STANDARD; PRT; 521 AA.
AC O00505; O00191; O43195; Q96AA7;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Importin alpha-3 subunit (Karyopherin alpha-3 subunit) (SRP1-gamma).
GN KPNA3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=97298718; PubMed=9154134;
RA Takeda S., Fujiwara T., Shimizu F., Kawai A., Shinomiya K.,
RA Okuno S., Ozaki K., Katagiri T., Shimada Y., Nagata M.,
RA Watanabe T., Takaichi A., Kuga Y., Suzuki M., Hishigaki H.,
RA Takahashi E., Shin S., Nakamura Y., Hirai Y.;
RT "Isolation and mapping of karyopherin alpha 3 (KPNA3), a human gene
RT that is highly homologous to genes encoding Xenopus importin, yeast
RT SRP1 and human RCH1.";

RL Cytogenet. Cell Genet. 76:87-93(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=98055463; PubMed=9395085;
 RA Koehler M., Ansieau S., Prehn S., Leutz A., Haller H., Hartmann E.;
 RT "Cloning of two novel human importin-alpha subunits and analysis of
 RT the expression pattern of the importin-alpha protein family.";
 RL FEBS Lett. 417:104-108(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98118555; PubMed=9435235;
 RA Nachury M.V., Ryder U.W., Lamond A.I., Weis K.;
 RT "Cloning and characterization of hSRP1 gamma, a tissue-specific
 RT nuclear transport factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:582-587(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Binds specifically and directly to substrates containing
 CC either a simple or bipartite NLS motif. Promotes docking of import
 CC substrates to the nuclear envelope. Seems to act as a cytosolic
 CC receptor for both simple and bipartite NLS motifs (By similarity).
 CC -!- SUBUNIT: Forms a complex with importin beta-1 subunit.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
 CC -!- TISSUE SPECIFICITY: UBIQUITOUS. HIGHEST LEVELS IN HEART AND
 CC SKELETAL MUSCLE.
 CC -!- SIMILARITY: Belongs to the importin alpha family.
 CC -!- SIMILARITY: Contains 8 ARM repeats.
 CC -!- CAUTION: Was known in Ref.2 as importin alpha-4.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; D89618; BAA20378.1; -.
DR EMBL; Y12394; CAA73026.1; -.
DR EMBL; AF034756; AAB87693.1; -.
DR EMBL; BC017355; AAH17355.1; -.
DR EMBL; BC024202; AAH24202.1; -.
DR HSSP; Q02821; 1BK5.
DR Genew; HGNC:6396; KPNA3.
DR MIM; 601892; -.
DR GO; GO:0005647; C:importin, alpha-subunit; TAS.
DR GO; GO:0005643; C:nuclear pore; TAS.
DR GO; GO:0008139; F:nuclear localization sequence binding; TAS.
DR GO; GO:0006607; P:NLS-bearing substrate-nucleus import; TAS.
DR GO; GO:0006461; P:protein complex assembly; TAS.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR000225; Armadillo.
DR InterPro; IPR002652; ImportinA_B.
DR Pfam; PF00514; Armadillo_seg; 8.
DR Pfam; PF01749; IBB; 1.
DR SMART; SM00185; ARM; 8.
DR PROSITE; PS50176; ARM_REPEAT; 3.
KW Transport; Protein transport; Repeat; Nuclear protein; Polymorphism.
FT DOMAIN 11 51 IBB.
FT DOMAIN 26 41 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT REPEAT 107 149 ARM 1.
FT REPEAT 150 194 ARM 2.
FT REPEAT 195 233 ARM 3.
FT REPEAT 234 278 ARM 4.
FT REPEAT 279 318 ARM 5.
FT REPEAT 319 360 ARM 6.
FT REPEAT 361 400 ARM 7.
FT REPEAT 401 443 ARM 8.
FT DOMAIN 444 521 NLS RECOGNITION REGION.
FT VARIANT 291 291 P -> S (in dbSNP:1043015).
FT /FTId=VAR_014454.
FT CONFLICT 34 34 V -> M (IN REF. 2).
FT CONFLICT 103 103 R -> Q (IN REF. 1).
FT CONFLICT 154 154 V -> G (IN REF. 3).
FT CONFLICT 236 236 P -> T (IN REF. 3).
FT CONFLICT 237 237 M -> L (IN REF. 2).
FT CONFLICT 256 256 I -> V (IN REF. 2).
FT CONFLICT 259 259 L -> V (IN REF. 2).
SQ SEQUENCE 521 AA; 57811 MW; C4FF132C3F346B7F CRC64;

Query Match 54.3%; Score 44; DB 1; Length 521;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 KFPPPPQETV 14
| ||||| |||
Db 231 KDPPPPMETV 240

RESULT 2
IMA3_MOUSE

ID IMA3_MOUSE STANDARD; PRT; 521 AA.
 AC 035344;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Importin alpha-3 subunit (Karyopherin alpha-3 subunit) (Importin alpha
 DE Q2).
 GN KPNA3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98034160; PubMed=9369227;
 RA Tsuji L., Takumi T., Imamoto N., Yoneda Y.;
 RT "Identification of novel homologues of mouse importin alpha, the
 RT alpha subunit of the nuclear pore-targeting complex, and their
 RT tissue-specific expression.";
 RL FEBS Lett. 416:30-34(1997).
 CC -!- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING
 CC EITHER A SIMPLE OR BIPARTITE NLS MOTIF. PROMOTES DOCKING OF IMPORT
 CC SUBSTRATES TO THE NUCLEAR PORE COMPLEX (NPC). THE COMPLEX IS
 CC SUBSEQUENTLY TRANSLOCATED THROUGH THE PORE BY AN ENERGY REQUIRING,
 CC RAN-DEPENDENT MECHANISM. AT THE NUCLEOPLASMIC SIDE OF THE NPC, THE
 CC THREE COMPONENTS SEPARATE AND IMPORTIN-ALPHA AND -BETA ARE RE-
 CC EXPORTED FROM THE NUCLEUS TO THE CYTOPLASM.
 CC -!- SUBUNIT: Forms a complex with importin beta-1 subunit.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
 CC -!- TISSUE SPECIFICITY: DETECTED MORE OR LESS IN ALL TISSUES EXAMINED
 CC (EHRlich ASCITES TUMOR CELLS, TESTIS, KIDNEY, SPLEEN, LIVER,
 CC HEART, LUNG, THYMUS, SKELETAL MUSCLE, CEREBELLUM AND BRAIN
 CC (WITHOUT CEREBELLUM)).
 CC -!- DOMAIN: CONSISTS OF AN N-TERMINAL HYDROPHILIC REGION, A
 CC HYDROPHOBIC CENTRAL REGION COMPOSED OF 8 REPEATS, AND A SHORT
 CC HYDROPHILIC C-TERMINUS. THE N-TERMINAL HYDROPHILIC REGION CONTAINS
 CC THE IMPORTIN BETA BINDING DOMAIN (IBB DOMAIN), WHICH IS SUFFICIENT
 CC FOR BINDING IMPORTIN BETA AND ESSENTIAL FOR NUCLEAR PROTEIN
 CC IMPORT.
 CC -!- SIMILARITY: Belongs to the importin alpha family.
 CC -!- SIMILARITY: Contains 8 ARM repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF020772; AAC53372.1; -.
 DR HSSP; Q02821; 1BK5.
 DR MGD; MGI:1100863; Kpna3.
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR000225; Armadillo.
 DR InterPro; IPR002652; ImportinA_B.
 DR Pfam; PF00514; Armadillo_seg; 8.

DR Pfam; PF01749; IBB; 1.
 DR SMART; SM00185; ARM; 8.
 DR PROSITE; PS50176; ARM_REPEAT; 3.
 KW Transport; Protein transport; Repeat; Nuclear protein.
 FT DOMAIN 11 51 IBB.
 FT DOMAIN 26 41 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT REPEAT 107 149 ARM 1.
 FT REPEAT 150 194 ARM 2.
 FT REPEAT 195 233 ARM 3.
 FT REPEAT 234 278 ARM 4.
 FT REPEAT 279 318 ARM 5.
 FT REPEAT 319 360 ARM 6.
 FT REPEAT 361 400 ARM 7.
 FT REPEAT 401 443 ARM 8.
 FT DOMAIN 444 521 NLS RECOGNITION REGION.
 FT DOMAIN 233 236 POLY-PRO.
 SQ SEQUENCE 521 AA; 57772 MW; 78501CE1AAB3A260 CRC64;

Query Match 54.3%; Score 44; DB 1; Length 521;
 Best Local Similarity 80.0%; Pred. No. 30;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 KFPPPPQETV 14
 | |||| |||
 Db 231 KDPPPPMETV 240

RESULT 3

SON_HUMAN

ID SON HUMAN STANDARD; PRT; 2426 AA.
 AC P18583; O14487; O95981; Q14120; Q9H7B1; Q9P070; Q9P072; Q9UKP9;
 AC Q9UPY0;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE SON protein (SON3) (Negative regulatory element-binding protein) (NRE-
 DE binding protein) (DBP-5) (Bax antagonist selected in saccharomyces 1)
 DE (BASS1).
 GN SON OR NREBP OR DBP5 OR C21ORF50 OR KIAA1019.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS A; B; C; D; E AND F).
 RX MEDLINE=21564202; PubMed=11707072;
 RA Reymond A., Friedli M., Neergaard Henrichsen C., Chapot F.,
 RA Deutsch S., UCLA C., Rossier C., Lyle R., Guipponi M.,
 RA Antonarakis S.E.;
 RT "From PREDs and open reading frames to cDNA isolation: revisiting the
 RT human chromosome 21 transcription map.";
 RL Genomics 78:46-54(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM G).
 RC TISSUE=Liver;
 RX MEDLINE=21316479; PubMed=11306577;
 RA Sun C.-T., Lo W.-Y., Wang I.-H., Lo Y.-H., Shiou S.-R., Lai C.-K.,

RA Ting L.-P.;
 RT "Transcription repression of human hepatitis B virus genes by negative
 RT regulatory element-binding protein/SON.";
 RL J. Biol. Chem. 276:24059-24067(2001).
 RN [3]
 RP SEQUENCE OF 1-689 FROM N.A. (ISOFORM H).
 RC TISSUE=Placenta;
 RA Casadei R., Strippoli P., D'Addabbo P., Canaider S., Lenzi L.,
 RA Vitale L., Giannone S., Carinci P., Zannotti M.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP SEQUENCE OF 1-130 FROM N.A.
 RC TISSUE=Smooth muscle;
 RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
 RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
 RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
 RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
 RA Isogai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
 RN [5]
 RP SEQUENCE OF 1-114 FROM N.A.
 RC TISSUE=Blood;
 RA Ye M., Zhang Q.H., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
 RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
 RT "Human partial CDS from cd34+ stem cells.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
 RN [6]
 RP SEQUENCE OF 437-2426 FROM N.A. (ISOFORM B).
 RC TISSUE=Brain;
 RX MEDLINE=99397452; PubMed=10470851;
 RA Kikuno R., Nagase T., Ishikawa K.-I., Hirose M., Miyajima N.,
 RA Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIV.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:197-205(1999).
 RN [7]
 RP SEQUENCE OF 554-2426 FROM N.A. (ISOFORM A).
 RX MEDLINE=92049296; PubMed=1944255;
 RA Chumakov I.M., Berdichevskii F.B., Sokolova N.V., Reznikov M.V.,
 RA Prasolov V.S.;
 RT "Identification of a protein product of a novel human gene SON and
 RT the biological effect upon administering a changed form of this gene
 RT into mammalian cells.";
 RL Mol. Biol. (Mosk) 25:731-740(1991).
 RN [8]
 RP SEQUENCE OF 709-1079 FROM N.A. (ISOFORM I).
 RC TISSUE=Placenta;
 RX MEDLINE=93062885; PubMed=1435774;
 RA Bliskovskii V.V., Kirillov A.V., Zakhariev V.M., Chumakov I.M.;
 RT "The human son gene: the large and small transcripts contains various
 RT 5'-terminal sequences.";
 RL Mol. Biol. (Mosk) 26:807-812(1992).
 RN [9]
 RP SEQUENCE OF 1009-1131 FROM N.A.
 RC TISSUE=Placenta;

RX MEDLINE=93062884; PubMed=1435773;
 RA Bliskovskii V.V., Berdichevskii F.B., Tkachenko A.V., Belova M.E.,
 RA Chumakov I.M.;
 RT "Coding part of the son gene small transcript contains four areas of
 RT complete tandem repeats.";
 RL Mol. Biol. (Mosk) 26:793-806(1992).
 RN [10]
 RP SEQUENCE OF 1145-2426 FROM N.A. (ISOFORM F).
 RX MEDLINE=93048367; PubMed=1424986;
 RA Mattioni T., Hume C.R., Konigorski S., Hayes P., Osterweil Z.,
 RA Lee J.S.;
 RT "A cDNA clone for a novel nuclear protein with DNA binding
 RT activity.";
 RL Chromosoma 101:618-624(1992).
 RN [11]
 RP SEQUENCE OF 1692-2175 FROM N.A. (ISOFORM A).
 RX MEDLINE=89039788; PubMed=3054499;
 RA Berdichevskii F.B., Chumakov I.M., Kiselev L.L.;
 RT "Decoding of the primary structure of the son3 region in human
 RT genome: identification of a new protein with unusual structure and
 RT homology with DNA-binding proteins.";
 RL Mol. Biol. (Mosk) 22:794-801(1988).
 RN [12]
 RP SEQUENCE OF 1939-2426 FROM N.A. (ISOFORM J).
 RC TISSUE=Cerebellum;
 RX MEDLINE=99439804; PubMed=10509013;
 RA Greenhalf W., Lee J., Chaudhuri B.;
 RT "A selection system for human apoptosis inhibitors using yeast.";
 RL Yeast 15:1307-1321(1999).
 CC -!- FUNCTION: Represses hepatitis B virus (HBV) core promoter activity
 CC and transcription of HBV genes and production of HBV virions.
 CC Binds to the consensus DNA sequence: 5'-GA[GT]AN[CG][AG]CC-3'.
 CC Might protect cells from apoptosis. Might be involved in pre-mRNA
 CC splicing (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear with a speckled distribution.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=10;
 CC Comment=Experimental confirmation may be lacking for some
 CC isoforms;
 CC Name=F;
 CC IsoId=P18583-1; Sequence=Displayed;
 CC Name=A;
 CC IsoId=P18583-2; Sequence=VSP_004401, VSP_004402, VSP_004403;
 CC Name=B;
 CC IsoId=P18583-3; Sequence=VSP_004404, VSP_004405;
 CC Name=C;
 CC IsoId=P18583-4; Sequence=VSP_004406, VSP_004407;
 CC Name=D;
 CC IsoId=P18583-5; Sequence=VSP_004403;
 CC Name=E;
 CC IsoId=P18583-6; Sequence=VSP_004408, VSP_004409;
 CC Name=G;
 CC IsoId=P18583-7; Sequence=VSP_004410;
 CC Name=H;
 CC IsoId=P18583-8; Sequence=VSP_004411, VSP_004412;
 CC Name=I;
 CC IsoId=P18583-9; Sequence=VSP_004413;

```

CC      Name=J;
CC      IsoId=P18583-10; Sequence=VSP_004414, VSP_004415;
CC      -!- TISSUE SPECIFICITY: Widely expressed, with the higher expression
CC          seen in leukocyte and heart.
CC      -!- DOMAIN: Contains 8 types of repeats which are distributed in 3
CC          regions.
CC      -!- MISCELLANEOUS: Colocalizes with the pre-mRNA splicing factor
CC          SFRS2/SC-35.
CC      -!- SIMILARITY: Contains 1 G-patch domain.
CC      -!- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.
CC      -!- CAUTION: Isoform A sequence from Ref.7 differs from that shown due
CC          to a frameshift.
CC      -!- CAUTION: Isoform F sequence from Ref.10 differs from that shown
CC          due to a frameshift.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF380179; AAL34497.1; -.
DR      EMBL; X63753; CAA45282.1; ALT_FRAME.
DR      EMBL; M36428; AAA36624.1; -.
DR      EMBL; AF380180; AAL34498.1; -.
DR      EMBL; AF380181; AAL34499.1; -.
DR      EMBL; AF380182; AAL34500.1; -.
DR      EMBL; AF380183; AAL34501.1; -.
DR      EMBL; AF380184; AAL34502.1; -.
DR      EMBL; AY026895; AAK07692.1; -.
DR      EMBL; AF435977; AAL30810.1; -.
DR      EMBL; X63751; CAC69885.1; -.
DR      EMBL; AB028942; BAA82971.1; -.
DR      EMBL; X63071; CAA44793.1; ALT_FRAME.
DR      EMBL; AF139897; AAD50078.1; -.
DR      EMBL; AK024752; BAB14985.1; -.
DR      EMBL; AF161428; AAF28988.1; -.
DR      EMBL; AF161430; AAF28990.1; -.
DR      Genew; HGNC:11183; SON.
DR      GK; P18583; -.
DR      MIM; 182465; -.
DR      GO; GO:0008189; F:apoptosis inhibitor activity; IDA.
DR      GO; GO:0003677; F:DNA binding; TAS.
DR      GO; GO:0006916; P:anti-apoptosis; IDA.
DR      InterPro; IPR001159; DS_RBD.
DR      InterPro; IPR000467; G_patch.
DR      Pfam; PF00035; dsrm; 1.
DR      Pfam; PF01585; G-patch; 1.
DR      SMART; SM00358; DSRM; 1.
DR      SMART; SM00443; G_patch; 1.
DR      PROSITE; PS50137; DS_RBD; 1.
DR      PROSITE; PS50174; G_PATCH; 1.
KW      RNA-binding; DNA-binding; Nuclear protein; Repeat;
KW      Alternative splicing.
FT      DOMAIN      726      895      17 X 10 AA TANDEM REPEATS OF L-A-[ST]-

```

FT

[NSG]-[TS]-MDSQM.

Query Match 54.3%; Score 44; DB 1; Length 2426;
Best Local Similarity 58.3%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EMKFPPPPQETV 14
| : ||||:|:
Db 1705 EKEVPPPPKETL 1716

RESULT 4

IMA4_HUMAN

ID IMA4_HUMAN STANDARD; PRT; 521 AA.
AC O00629; O00190;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Importin alpha-4 subunit (Karyopherin alpha-4 subunit) (Qip1 protein).
GN KPNA4 OR QIP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97312517; PubMed=9168958;
RA Seki T., Tada S., Katada T., Enomoto T.;
RT "Cloning of a cDNA encoding a novel importin-alpha homologue, Qip1:
RT discrimination of Qip1 and Rchl from hSrpl by their ability to
RT interact with DNA helicase Q1/RecQL.";
RL Biochem. Biophys. Res. Commun. 234:48-53(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=98055463; PubMed=9395085;
RA Koehler M., Ansieau S., Prehn S., Leutz A., Haller H., Hartmann E.;
RT "Cloning of two novel human importin-alpha subunits and analysis of
RT the expression pattern of the importin-alpha protein family.";
RL FEBS Lett. 417:104-108(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING
 CC EITHER A SIMPLE OR BIPARTITE NLS MOTIF. PROMOTES DOCKING OF IMPORT
 CC SUBSTRATES TO THE NUCLEAR PORE COMPLEX (NPC). THE COMPLEX IS
 CC SUBSEQUENTLY TRANSLOCATED THROUGH THE PORE BY AN ENERGY REQUIRING,
 CC RAN-DEPENDENT MECHANISM. AT THE NUCLEOPLASMIC SIDE OF THE NPC, THE
 CC THREE COMPONENTS SEPARATE AND IMPORTIN-ALPHA AND -BETA ARE RE-
 CC EXPORTED FROM THE NUCLEUS TO THE CYTOPLASM (BY SIMILARITY).
 CC -!- SUBUNIT: Forms a complex with importin beta-1 subunit.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
 CC -!- TISSUE SPECIFICITY: Highly expressed in testis, ovary, small
 CC intestine, heart, skeletal muscle, lung and pancreas, but barely
 CC detectable in kidney, thymus, colon and peripheral blood
 CC leukocytes.
 CC -!- DOMAIN: CONSISTS OF AN N-TERMINAL HYDROPHILIC REGION, A
 CC HYDROPHOBIC CENTRAL REGION COMPOSED OF 8 REPEATS, AND A SHORT
 CC HYDROPHILIC C-TERMINUS. THE N-TERMINAL HYDROPHILIC REGION CONTAINS
 CC THE IMPORTIN BETA BINDING DOMAIN (IBB DOMAIN), WHICH IS SUFFICIENT
 CC FOR BINDING IMPORTIN BETA AND ESSENTIAL FOR NUCLEAR PROTEIN
 CC IMPORT.
 CC -!- SIMILARITY: Belongs to the importin alpha family.
 CC -!- SIMILARITY: Contains 8 ARM repeats.
 CC -!- CAUTION: Was known in Ref.2 as importin alpha-3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB002533; BAA19546.1; -.
 DR EMBL; U93240; AAC25605.1; -.
 DR EMBL; Y12393; CAA73025.1; -.
 DR EMBL; BC028691; AAH28691.1; -.
 DR EMBL; BC034493; AAH34493.1; -.
 DR PIR; JC5505; JC5505.
 DR HSSP; Q02821; 1BK5.
 DR Genew; HGNC:6397; KPNA4.
 DR MIM; 602970; -.
 DR GO; GO:0005647; C:importin, alpha-subunit; TAS.
 DR GO; GO:0006607; P:NLS-bearing substrate-nucleus import; TAS.
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR000225; Armadillo.
 DR InterPro; IPR002652; ImportinA_B.
 DR Pfam; PF00514; Armadillo_seg; 8.
 DR Pfam; PF01749; IBB; 1.
 DR SMART; SM00185; ARM; 8.
 DR PROSITE; PS50176; ARM_REPEAT; 3.
 KW Transport; Protein transport; Repeat; Nuclear protein.

FT	DOMAIN	11	51	IBB.
FT	DOMAIN	26	41	NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT	REPEAT	107	149	ARM 1.
FT	REPEAT	150	194	ARM 2.
FT	REPEAT	195	233	ARM 3.
FT	REPEAT	234	278	ARM 4.
FT	REPEAT	279	318	ARM 5.
FT	REPEAT	319	360	ARM 6.
FT	REPEAT	361	400	ARM 7.
FT	REPEAT	401	443	ARM 8.
FT	DOMAIN	444	521	NLS RECOGNITION REGION.
FT	DOMAIN	233	236	POLY-PRO.
FT	CONFLICT	241	241	Q -> T (IN REF. 2; CAA73025).
SQ	SEQUENCE	521 AA;	57886 MW;	D98BC45002C9F57E CRC64;

Query Match 53.1%; Score 43; DB 1; Length 521;
 Best Local Similarity 70.0%; Pred. No. 42;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 KFPPPPQETV 14
 | |||| ||:
 Db 231 KDPPPPMETI 240

RESULT 5

IMA4_MOUSE

ID IMA4_MOUSE STANDARD; PRT; 521 AA.
 AC O35343;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Importin alpha-4 subunit (Karyopherin alpha-4 subunit) (Importin alpha Q1).
 DE
 GN KPNA4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98034160; PubMed=9369227;
 RA Tsuji L., Takumi T., Imamoto N., Yoneda Y.;
 RT "Identification of novel homologues of mouse importin alpha, the
 RT alpha subunit of the nuclear pore-targeting complex, and their
 RT tissue-specific expression."
 RL FEBS Lett. 416:30-34(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Limb;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING
 CC EITHER A SIMPLE OR BIPARTITE NLS MOTIF. PROMOTES DOCKING OF IMPORT
 CC SUBSTRATES TO THE NUCLEAR PORE COMPLEX (NPC). THE COMPLEX IS
 CC SUBSEQUENTLY TRANSLOCATED THROUGH THE PORE BY AN ENERGY REQUIRING,
 CC RAN-DEPENDENT MECHANISM. AT THE NUCLEOPLASMIC SIDE OF THE NPC, THE
 CC THREE COMPONENTS SEPARATE AND IMPORTIN-ALPHA AND -BETA ARE RE-
 CC EXPORTED FROM THE NUCLEUS TO THE CYTOPLASM.
 CC -!- SUBUNIT: Forms a complex with importin beta-1 subunit.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
 CC -!- TISSUE SPECIFICITY: DETECTED MORE OR LESS IN ALL TISSUES EXAMINED
 CC (EHRlich ASCITES TUMOR CELLS, TESTIS, KIDNEY, SPLEEN, LIVER,
 CC HEART, LUNG, THYMUS, SKELETAL MUSCLE, CEREBELLUM AND BRAIN
 CC (WITHOUT CEREBELLUM)). MULTIPLE-SIZED TRANSCRIPTS WERE HIGHLY
 CC EXPRESSED, ESPECIALLY IN TESTIS.
 CC -!- DOMAIN: CONSISTS OF AN N-TERMINAL HYDROPHILIC REGION, A
 CC HYDROPHOBIC CENTRAL REGION COMPOSED OF 8 REPEATS, AND A SHORT
 CC HYDROPHILIC C-TERMINUS. THE N-TERMINAL HYDROPHILIC REGION CONTAINS
 CC THE IMPORTIN BETA BINDING DOMAIN (IBB DOMAIN), WHICH IS SUFFICIENT
 CC FOR BINDING IMPORTIN BETA AND ESSENTIAL FOR NUCLEAR PROTEIN
 CC IMPORT.
 CC -!- SIMILARITY: Belongs to the importin alpha family.
 CC -!- SIMILARITY: Contains 8 ARM repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF020771; AAC53371.1; -.
 DR EMBL; BC052162; AAH52162.1; -.
 DR HSSP; Q02821; 1BK5.
 DR MGD; MGI:1100848; Kpna4.
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR000225; Armadillo.
 DR InterPro; IPR002652; ImportinA_B.
 DR Pfam; PF00514; Armadillo_seg; 8.
 DR Pfam; PF01749; IBB; 1.
 DR SMART; SM00185; ARM; 7.
 DR PROSITE; PS50176; ARM_REPEAT; 3.
 KW Transport; Protein transport; Repeat; Nuclear protein.
 FT DOMAIN 11 51 IBB.

FT	DOMAIN	26	41	NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT	REPEAT	107	149	ARM 1.
FT	REPEAT	150	194	ARM 2.
FT	REPEAT	195	233	ARM 3.
FT	REPEAT	234	278	ARM 4.
FT	REPEAT	279	318	ARM 5.
FT	REPEAT	319	360	ARM 6.
FT	REPEAT	361	400	ARM 7.
FT	REPEAT	401	443	ARM 8.
FT	DOMAIN	444	521	NLS RECOGNITION REGION.
FT	DOMAIN	233	236	POLY-PRO.
SQ	SEQUENCE	521 AA; 57922 MW; A0578EC9D411E060 CRC64;		

Query Match 53.1%; Score 43; DB 1; Length 521;
 Best Local Similarity 70.0%; Pred. No. 42;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 KFPPPPQETV 14
 | | | | | | | :
 Db 231 KDPPPPMETI 240

RESULT 6

ABCI_SCHPO

ID ABCI_SCHPO STANDARD; PRT; 610 AA.
 AC Q92338;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ABC1 protein homolog, mitochondrial precursor.
 GN ABC1 OR SPBC2D10.18.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96242153; PubMed=8668131;
 RA Bonnefoy N., Kermorgant M., Brivet-Chevillotte P., Dujardin G.;
 RT "Cloning by functional complementation, and inactivation, of the
 RT Schizosaccharomyces pombe homologue of the Saccharomyces cerevisiae
 RT gene ABC1.";
 RL Mol. Gen. Genet. 251:204-210(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of *Schizosaccharomyces pombe*.";
 RL Nature 415:871-880(2002).

CC -!- SUBCELLULAR LOCATION: Mitochondrial (Probable).

CC -!- SIMILARITY: Belongs to the ABC1 family.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X91616; CAA62818.1; -.

DR EMBL; AL031788; CAA21176.1; -.

DR PIR; S71111; S71110.

DR GeneDB_SPombe; SPBC2D10.18; -.

DR InterPro; IPR004147; ABC1.

DR Pfam; PF03109; ABC1; 1.

KW Chaperone; Mitochondrion; Transit peptide.

FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).

FT CHAIN ? 610 ABC1 PROTEIN HOMOLOG.

SQ SEQUENCE 610 AA; 68566 MW; ACFED14EEEE8353D CRC64;

Query Match 53.1%; Score 43; DB 1; Length 610;

Best Local Similarity 46.2%; Pred. No. 50;

Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13

:|::: |||:|

Db 562 MLDLRLQPPPEET 574

RESULT 7

FATH_HUMAN

ID FATH_HUMAN STANDARD; PRT; 4590 AA.

AC Q14517;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Cadherin-related tumor suppressor homolog precursor (Fat protein

DE homolog).

GN FAT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymphocytes;
 RX MEDLINE=96163873; PubMed=8586420;
 RA Dunne J., Hanby A.M., Poulsom R., Jones T.A., Sheer D., Chin W.G.,
 RA Da S.M., Zhao Q., Beverley P.C.L., Owen M.J.;
 RT "Molecular cloning and tissue expression of FAT, the human homologue
 RT of the Drosophila fat gene that is located on chromosome 4q34-q35 and
 RT encodes a putative adhesion molecule.";
 RL Genomics 30:207-223(1995).
 CC -!- FUNCTION: Could function as a cell-adhesion protein.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in many epithelial and some
 CC endothelial and smooth muscle cells.
 CC -!- SIMILARITY: Contains 34 cadherin domains.
 CC -!- SIMILARITY: Contains 5 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 laminin G-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X87241; CAA60685.1; -.
 DR HSSP; P01132; 1EGF.
 DR Genew; HGNC:3595; FAT.
 DR MIM; 600976; -.
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR GO; GO:0007267; P:cell-cell signaling; TAS.
 DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
 DR InterPro; IPR000152; Asx_hydroxyl_S.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR008985; ConA_like_lec_gl.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR001791; Laminin_G.
 DR Pfam; PF00028; cadherin; 33.
 DR Pfam; PF00008; EGF; 5.
 DR Pfam; PF00054; laminin_G; 1.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 31.
 DR SMART; SM00179; EGF_CA; 1.
 DR SMART; SM00282; LamG; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS00232; CADHERIN_1; 16.
 DR PROSITE; PS50268; CADHERIN_2; 33.
 DR PROSITE; PS00022; EGF_1; 4.

DR	PROSITE; PS01186; EGF_2; 1.			
DR	PROSITE; PS50026; EGF_3; 5.			
DR	PROSITE; PS01187; EGF_CA; 1.			
DR	PROSITE; PS50025; LAM_G_DOMAIN; 1.			
KW	Cell adhesion; Signal; Glycoprotein; Transmembrane; Calcium-binding;			
KW	Repeat; EGF-like domain.			
FT	SIGNAL	1	21	POTENTIAL.
FT	CHAIN	22	4590	CADHERIN-RELATED TUMOR SUPPRESSOR
FT				HOMOLOG.
FT	DOMAIN	22	4183	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	4184	4204	POTENTIAL.
FT	DOMAIN	4205	4590	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	22	149	CADHERIN 1.
FT	DOMAIN	150	256	CADHERIN 2.
FT	DOMAIN	257	361	CADHERIN 3.
FT	DOMAIN	362	463	CADHERIN 4.
FT	DOMAIN	464	569	CADHERIN 5.
FT	DOMAIN	570	716	CADHERIN 6.
FT	DOMAIN	717	822	CADHERIN 7.
FT	DOMAIN	823	927	CADHERIN 8.
FT	DOMAIN	928	1034	CADHERIN 9.
FT	DOMAIN	1035	1138	CADHERIN 10.
FT	DOMAIN	1139	1245	CADHERIN 11.
FT	DOMAIN	1246	1345	CADHERIN 12.
FT	DOMAIN	1346	1456	CADHERIN 13.
FT	DOMAIN	1457	1562	CADHERIN 14.
FT	DOMAIN	1563	1670	CADHERIN 15.
FT	DOMAIN	1671	1769	CADHERIN 16.
FT	DOMAIN	1770	1882	CADHERIN 17.
FT	DOMAIN	1883	1982	CADHERIN 18.
FT	DOMAIN	1983	2084	CADHERIN 19.
FT	DOMAIN	2085	2185	CADHERIN 20.
FT	DOMAIN	2186	2286	CADHERIN 21.
FT	DOMAIN	2287	2393	CADHERIN 22.
FT	DOMAIN	2394	2495	CADHERIN 23.
FT	DOMAIN	2496	2599	CADHERIN 24.
FT	DOMAIN	2600	2705	CADHERIN 25.
FT	DOMAIN	2706	2811	CADHERIN 26.
FT	DOMAIN	2812	2920	CADHERIN 27.
FT	DOMAIN	2921	3024	CADHERIN 28.
FT	DOMAIN	3025	3127	CADHERIN 29.
FT	DOMAIN	3128	3232	CADHERIN 30.
FT	DOMAIN	3233	3337	CADHERIN 31.
FT	DOMAIN	3338	3442	CADHERIN 32.
FT	DOMAIN	3443	3546	CADHERIN 33.
FT	DOMAIN	3547	3649	CADHERIN 34.
FT	DOMAIN	3790	3827	EGF-LIKE 1.
FT	DOMAIN	3831	4011	LAMININ G-LIKE.
FT	DOMAIN	4013	4051	EGF-LIKE 2.
FT	DOMAIN	4052	4089	EGF-LIKE 3.
FT	DOMAIN	4090	4126	EGF-LIKE 4.
FT	DOMAIN	4127	4163	EGF-LIKE 5.
FT	CARBOHYD	40	40	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	333	333	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	660	660	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	740	740	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	791	791	N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	CARBOHYD	998	998	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1426	1426	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1551	1551	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1751	1751	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1867	1867	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1905	1905	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1943	1943	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1994	1994	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2328	2328	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2467	2467	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3326	3326	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3424	3424	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3446	3446	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3615	3615	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3642	3642	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3718	3718	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	4154	4154	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	4590 AA;	506273 MW;	04483CCD000E60A7 CRC64;

Query Match 53.1%; Score 43; DB 1; Length 4590;
 Best Local Similarity 54.5%; Pred. No. 4.2e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEMKFPPPPQE 12
 :| |||||::
 Db 4433 IESDFPPPPED 4443

RESULT 8

CRN1_MOUSE

ID CRN1_MOUSE STANDARD; PRT; 690 AA.
 AC Q9CQC1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Crooked neck-like protein 1 (Crooked neck homolog).
 GN CRNKL1.
 OS Mus musculus (Mouse), and
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090, 10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse; STRAIN=C57BL/6J; TISSUE=Embryo, and Lung;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Rat; STRAIN=Sprague-Dawley; TISSUE=Brain;
 RA Amada N., Araki K., Nawa H.;
 RT "Cloning and characterization of the crooked neck gene.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: Involved in pre-mRNA splicing process (By similarity).
 CC -!- SUBUNIT: Present in a spliceosome complex assembled in vitro
 CC containing CRNKL1, HPRP8BP and SNRPB2 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear; in nuclear speckles (By
 CC similarity).
 CC -!- SIMILARITY: Contains 16 HAT repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AK012962; BAB28572.1; -.
 DR EMBL; AK004749; BAB23530.1; -.
 DR EMBL; BC029187; AAH29187.1; -.

DR EMBL; AF245018; AAK27972.1; -.
 DR MGD; MGI:1914127; Crnkl1.
 DR InterPro; IPR003107; HAT.
 DR InterPro; IPR008940; Prenyl_trans.
 DR InterPro; IPR008941; TPR-like.
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF02184; HAT; 12.
 KW Nuclear protein; mRNA processing; mRNA splicing; Repeat.
 FT REPEAT 61 93 HAT 1.
 FT REPEAT 95 127 HAT 2.
 FT REPEAT 129 161 HAT 3.
 FT REPEAT 163 194 HAT 4.
 FT REPEAT 196 227 HAT 5.
 FT REPEAT 229 264 HAT 6.
 FT REPEAT 266 300 HAT 7.
 FT REPEAT 310 342 HAT 8.
 FT REPEAT 344 378 HAT 9.
 FT REPEAT 388 424 HAT 10.
 FT REPEAT 459 491 HAT 11.
 FT REPEAT 493 527 HAT 12.
 FT REPEAT 529 560 HAT 13.
 FT REPEAT 565 606 HAT 14.
 FT REPEAT 608 646 HAT 15.
 FT REPEAT 648 673 HAT 16.
 FT DOMAIN 46 49 POLY-PRO.
 SQ SEQUENCE 690 AA; 83415 MW; 9FE84D7F1205B974 CRC64;

Query Match 52.5%; Score 42.5; DB 1; Length 690;
 Best Local Similarity 57.1%; Pred. No. 67;
 Matches 8; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy 2 LEMKFPPPPQETVT 15
 ||: ||||: :|
 Db 42 LEL-LPPPPQKIT 54

RESULT 9

CRN1_HUMAN

ID CRN1_HUMAN STANDARD; PRT; 848 AA.
 AC Q9BZJ0; Q8WYI5; Q9BZI9; Q9BZJ1; Q9BZJ2; Q9GZW7; Q9H8F8; Q9NQH5;
 AC Q9NYD8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Crooked neck-like protein 1 (Crooked neck homolog) (hCrn) (CGI-201)
 DE (MSTP021).
 GN CRNKL1 OR CRN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5).
 RX MEDLINE=21240337; PubMed=11342225;
 RA Lai C.-H., Chiu J.-Y., Lin W.-C.;
 RT "Identification of the human crooked neck gene by comparative gene
 RT identification.";

RL Biochim. Biophys. Acta 1517:449-454(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), CHARACTERIZATION, AND
 RP IDENTIFICATION IN SPLICEOSOMAL COMPLEX WITH HPRP8BP AND SNRPB2.
 RC TISSUE=T lymphoblast;
 RX MEDLINE=22079017; PubMed=12084575;
 RA Chung S., Zhou Z., Huddleston K.A., Harrison D.A., Reed R.,
 RA Coleman T.A., Rymond B.C.;
 RT "Crooked neck is a component of the human spliceosome and implicated
 RT in the splicing process.";
 RL Biochim. Biophys. Acta 1576:287-297(2002).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Heart;
 RA Xu Y.Y., Sun L.Z., Wu Q.Y., Liu Y.Q., Liu B., Zhao B., Wang X.Y.,
 RA Song L., Ye J., Sheng H., Gao Y., Zhang C.L., Zhang J., Wei Y.J.,
 RA Sun Y.H., Jiang Y.X., Zhao X.W., Liu S., Liu L.S., Ding J.F.,
 RA Gao R.L., Qiang B.Q., Yuan J.G., Liew C.C., Zhao M.S., Hui R.T.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehvaslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [5]
 RP SEQUENCE OF 220-848 FROM N.A. (ISOFORM 1/2/3).
 RC TISSUE=Placenta;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;

```

RT  "NEDO human cDNA sequencing project.";
RL  Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC  -!- FUNCTION: Involved in pre-mRNA splicing process.
CC  -!- SUBUNIT: Present in a spliceosome complex assembled in vitro
CC      containing CRNKL1, HPRP8BP and SNRPB2. Isoform 2 seems to be
CC      predominant in the spliceosome complex.
CC  -!- SUBCELLULAR LOCATION: Nuclear; in nuclear speckles. Colocalizes
CC      with core spliceosomal snRNP proteins. Also diffusely expressed in
CC      cytoplasm.
CC  -!- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=5;
CC      Name=1; Synonyms=type-II;
CC      IsoId=Q9BZJ0-1; Sequence=Displayed;
CC      Name=2; Synonyms=short;
CC      IsoId=Q9BZJ0-2; Sequence=VSP_002058;
CC      Name=3;
CC      IsoId=Q9BZJ0-3; Sequence=VSP_002059;
CC      Name=4; Synonyms=type-III;
CC      IsoId=Q9BZJ0-4; Sequence=VSP_002059, VSP_002062;
CC      Name=5; Synonyms=type-IV;
CC      IsoId=Q9BZJ0-5; Sequence=VSP_002060, VSP_002061;
CC  -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in testis.
CC      Not expressed in brain and lung.
CC  -!- SIMILARITY: Contains 17 HAT repeats.
CC  -!- CAUTION: Ref.4 sequence differs from that shown due to erroneous
CC      gene model prediction.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF255443; AAF65571.2; -.
DR  EMBL; AF318302; AAK01924.1; -.
DR  EMBL; AF318303; AAK01925.1; -.
DR  EMBL; AF318304; AAK01926.1; -.
DR  EMBL; AF318305; AAK01927.1; -.
DR  EMBL; AF111802; AAL39004.1; -.
DR  EMBL; AL035454; CAC01671.1; ALT_SEQ.
DR  EMBL; AK022908; BAB14303.1; ALT_INIT.
DR  EMBL; AK023246; BAB14485.1; ALT_INIT.
DR  EMBL; AK023728; BAB14659.1; -.
DR  Genew; HGNC:15762; CRNKL1.
DR  GK; Q9BZJ0; -.
DR  InterPro; IPR003107; HAT.
DR  InterPro; IPR008940; Prenyl_trans.
DR  InterPro; IPR008941; TPR-like.
DR  InterPro; IPR001440; TPR.
DR  Pfam; PF02184; HAT; 12.
DR  SMART; SM00386; HAT; 13.
KW  Nuclear protein; mRNA processing; mRNA splicing; Repeat;
KW  Alternative splicing.
FT  REPEAT      222      254      HAT 1.
FT  REPEAT      256      288      HAT 2.

```

FT	REPEAT	290	322	HAT 3.
FT	REPEAT	324	355	HAT 4.
FT	REPEAT	357	388	HAT 5.
FT	REPEAT	390	425	HAT 6.
FT	REPEAT	427	461	HAT 7.
FT	REPEAT	471	503	HAT 8.
FT	REPEAT	505	539	HAT 9.
FT	REPEAT	549	585	HAT 10.
FT	REPEAT	587	618	HAT 11.
FT	REPEAT	620	652	HAT 12.
FT	REPEAT	654	688	HAT 13.
FT	REPEAT	690	721	HAT 14.
FT	REPEAT	726	767	HAT 15.
FT	REPEAT	769	807	HAT 16.
FT	REPEAT	809	834	HAT 17.
FT	DOMAIN	207	210	POLY-PRO.
FT	VARSPLIC	1	161	Missing (in isoform 2).
FT				/FTId=VSP_002058.
FT	VARSPLIC	23	34	Missing (in isoform 3 and isoform 4).
FT				/FTId=VSP_002059.
FT	VARSPLIC	23	74	TSRLELRYSYSLARRHGSTEPLVLAWSSQFRRLTWGCALDAL
FT				HRSPCVAASQH -> ILASLLVSTALPTSSAAPGRRTPRAA
FT				ARRTRSLVTMETVPPPSRLKREVKGQ (in isoform
FT				5).
FT				/FTId=VSP_002060.
FT	VARSPLIC	75	848	Missing (in isoform 5).
FT				/FTId=VSP_002061.
FT	VARSPLIC	117	848	Missing (in isoform 4).
FT				/FTId=VSP_002062.
FT	CONFLICT	213	213	R -> K (IN REF. 2 AND 3).
FT	CONFLICT	513	513	V -> A (IN REF. 5; BAB14659).
SQ	SEQUENCE	848 AA;	100578 MW;	7C43D73D4EA5F834 CRC64;

Query Match 52.5%; Score 42.5; DB 1; Length 848;
 Best Local Similarity 57.1%; Pred. No. 83;
 Matches 8; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy 2 LEMKFPPPPQETVT 15
 ||: ||||: :|
 Db 203 LEL-LPPPPQQRIT 215

RESULT 10

FTN_DROME

ID FTN_DROME STANDARD; PRT; 182 AA.
 AC P35554; Q9VW65;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Flightin (Muscle protein 27).
 GN FLN OR FTN OR MP27 OR CG7445.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Flight muscle;
 RX MEDLINE=93252994; PubMed=8486738;
 RA Vigoreaux J.O., Saide J.D., Valgeirsdottir K., Pardue M.L.;
 RT "Flightin, a novel myofibrillar protein of *Drosophila*
 RT stretch-activated muscles.";
 RL J. Cell Biol. 121:587-598(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley; TISSUE=Ovary;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celniker S.E.;
 RT "A *Drosophila* full-length cDNA resource.";

RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 CC -!- FUNCTION: Possibly involved in the regulation of flight muscles
 CC contraction, possibly by modulating actin-myosin interaction.
 CC -!- TISSUE SPECIFICITY: Found only in indirect flight muscles (IFM).
 CC -!- DEVELOPMENTAL STAGE: Expressed in late pupal and adult stages.
 CC -!- PTM: Several forms of flightin are thought to be produced through
 CC post-translational modifications, possibly by phosphorylation.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z18858; CAA79309.1; -.
 DR EMBL; AE003515; AAF49084.1; -.
 DR EMBL; AY060802; AAL28350.1; -.
 DR PIR; A46436; A46436.
 DR FlyBase; FBgn0005633; fln.
 DR GO; GO:0005863; C:striated muscle thick filament; IDA.
 KW Muscle protein; Phosphorylation.
 FT DOMAIN 1 65 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 66 70 POLY-PRO.
 FT DOMAIN 71 182 ARG/LYS-RICH (BASIC).
 SQ SEQUENCE 182 AA; 20656 MW; 4855356B0FD24F1E CRC64;

Query Match 51.9%; Score 42; DB 1; Length 182;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQE 12
 | |:| |||| |
 Db 60 VEEIKAPPPPE 71

RESULT 11

CD20_MOUSE

ID CD20_MOUSE STANDARD; PRT; 291 AA.
 AC P19437;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE B-cell surface protein CD20 homolog (B-cell differentiation antigen
 DE LY-44).
 GN MS4A1 OR CD20 OR LY-44 OR MS4A2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89067519; PubMed=2461992;
 RA Tedder T.F., Klejman G., Distech C.M., Adler D.A., Schlossman S.F.,
 RA Saito H.;
 RT "Cloning of a complementary DNA encoding a new mouse B lymphocyte

RT differentiation antigen, homologous to the human B1 (CD20) antigen,
RT and localization of the gene to chromosome 19.";
RL J. Immunol. 141:4388-4394(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF
CC B-CELL ACTIVATION AND PROLIFERATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: Phosphorylated (By similarity).
CC -!- SIMILARITY: Belongs to the MS4A family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62541; AAA37394.1; -.
DR EMBL; AK017903; BAB30996.1; -.
DR PIR; A30558; A30558.
DR MGD; MGI:88321; Ms4a1.
DR InterPro; IPR007237; CD20.
DR Pfam; PF04103; CD20; 1.
KW B-cell activation; Antigen; Transmembrane; Phosphorylation;
KW Multigene family.
FT DOMAIN 1 44 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 45 65 POTENTIAL.
FT TRANSMEM 69 89 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 183 203 POTENTIAL.
FT DOMAIN 204 291 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 291 AA; 31958 MW; DF478ECD2C5C16FC CRC64;

Query Match 51.9%; Score 42; DB 1; Length 291;
Best Local Similarity 70.0%; Pred. No. 32;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EMKFPPPPQE 12
|: || ||||
Db 270 EINFAPPQE 279

RESULT 12

RHB4_HUMAN

ID RHB4_HUMAN STANDARD; PRT; 404 AA.

AC P58872;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Rhomboid-related protein 4 (EC 3.4.21.-) (Ventrhoid transmembrane protein).

GN RHBDL4 OR VRHO.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Fetal brain;

RX MEDLINE=21898529; PubMed=11900977;

RA Jasza J., Brand M.;

RT "Cloning and expression of Ventrhoid, a novel vertebrate homologue of the Drosophila EGF pathway gene Rhomboid.";

RL Mech. Dev. 113:73-77(2002).

CC -!- FUNCTION: May be involved in regulated intramembrane proteolysis and the subsequent release of functional polypeptides from their membrane anchors (By similarity).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

CC -!- SIMILARITY: Belongs to peptidase family S54.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AJ313480; CAC86145.1; -.

DR Genew; HGNC:19244; RHBDL4.

DR MEROPS; S54.950; -.

DR InterPro; IPR002048; EF-hand.

DR InterPro; IPR002610; Peptidase_S54.

DR Pfam; PF00036; efhand; 2.

DR Pfam; PF01694; Rhomboid; 1.

KW Hydrolase; Protease; Serine protease; Transmembrane.

FT TRANSMEM 164 184 POTENTIAL.

FT TRANSMEM 218 238 POTENTIAL.

FT TRANSMEM 250 270 POTENTIAL.

FT TRANSMEM 274 294 POTENTIAL.

FT TRANSMEM 303 325 POTENTIAL.
 FT TRANSMEM 338 358 POTENTIAL.
 FT TRANSMEM 371 391 POTENTIAL.
 FT ACT_SITE 230 230 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 278 278 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 343 343 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 404 AA; 45244 MW; BAE7319D4E2AB5AC CRC64;

Query Match 51.9%; Score 42; DB 1; Length 404;
 Best Local Similarity 60.0%; Pred. No. 45;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPP 10
 :|::| ||||
 Db 395 LLDLKLPPPP 404

RESULT 13

LEGU_CANEN

ID LEGU_CANEN STANDARD; PRT; 475 AA.
 AC P49046;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Legumain precursor (EC 3.4.22.34) (Asparaginyl endopeptidase).
 OS Canavalia ensiformis (Jack bean) (Horse bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
 OX NCBI_TaxID=3823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Seed;
 RX MEDLINE=95155263; PubMed=7852272;
 RA Takeda O., Miura Y., Mitta M., Matsushita H., Kato I., Abe Y.,
 RA Yokosawa H., Ishii S.;
 RT "Isolation and analysis of cDNA encoding a precursor of Canavalia
 RT ensiformis asparaginyl endopeptidase (legumain).";
 RL J. Biochem. 116:541-546(1994).
 RN [2]
 RP SEQUENCE OF 36-60, AND CHARACTERIZATION.
 RX MEDLINE=93155205; PubMed=8429028;
 RA Abe Y., Shirane K., Yokosawa H., Matsushita H., Mitta M.,
 RA Kato I., Ishii S.;
 RT "Asparaginyl endopeptidase of jack bean seeds. Purification,
 RT characterization, and high utility in protein sequence analysis.";
 RL J. Biol. Chem. 268:3525-3529(1993).
 RN [3]
 RP REVIEW.
 RX MEDLINE=95147717; PubMed=7845236;
 RA Ishii S.;
 RT "Legumain: asparaginyl endopeptidase.";
 RL Meth. Enzymol. 244:604-615(1994).
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins and small-molecule
 CC substrates at -Asn-|-Xaa- bonds.
 CC -!- SIMILARITY: Belongs to peptidase family C13.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; D31787; BAA06596.1; -.
 DR PIR; JX0344; JX0344.
 DR MEROPS; C13.001; -.
 DR InterPro; IPR001096; Peptidase_C13.
 DR Pfam; PF01650; Peptidase_C13; 1.
 DR PRINTS; PR00776; HEMOGLOBINASE.
 KW Hydrolase; Thiol protease; Signal; Zymogen.
 FT SIGNAL 1 15 POTENTIAL.
 FT PROPEP 16 35
 FT CHAIN 36 475 LEGUMAIN.
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT ACT_SITE 158 158 POTENTIAL.
 FT ACT_SITE 200 200 POTENTIAL.
 SQ SEQUENCE 475 AA; 52763 MW; 6D1C1D6872C5504C CRC64;

Query Match 51.9%; Score 42; DB 1; Length 475;
 Best Local Similarity 77.8%; Pred. No. 53;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 PPPPQETVT 15
 ||||:| ||
 Db 238 PPPPEEYVT 246

RESULT 14

MOC3_HUMAN

ID MOC3_HUMAN STANDARD; PRT; 460 AA.
 AC O95396;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Molybdenum cofactor synthesis protein 3 (Molybdopterin synthase
 DE sulfurylase) (MPT synthase sulfurylase).
 GN MOCS3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stallmeyer B., Coyne K.E., Wuebbens M.M., Johnson J.L.,
 RA Rajagopalan K.V., Mendel R.R.;
 RT "The cDNA sequence of MOCS3, human molybdopterin synthase
 RT sulfurylase.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehvaeslaiho M.H., Lerversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20";
 RL Nature 414:865-871(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Uterus;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Activates MPT synthase by the ATP dependant adenylation
 CC of its C-terminal residue (By similarity).
 CC -!- PATHWAY: Molybdenum cofactor biosynthesis.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC HESA/MOEB/THIF FAMILY.
 CC -!- SIMILARITY: Contains 1 rhodanese domain.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF102544; AAC72412.1; -.
 DR EMBL; AL034553; CAB53750.1; -.
 DR EMBL; BC015939; AAH15939.1; -.
 DR Genew; HGNC:15765; MOCS3.
 DR InterPro; IPR009036; MoeB.
 DR InterPro; IPR007901; MoeZ_MoeB.
 DR InterPro; IPR000205; NAD_BS.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000594; ThiF_domain.
 DR Pfam; PF05237; MoeZ_MoeB; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR Pfam; PF00899; ThiF; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANESE_3; 1.
 KW Ligase; Molybdenum cofactor biosynthesis.
 FT DOMAIN 347 458 RHODANESE.
 SQ SEQUENCE 460 AA; 49669 MW; 299A4E755173E324 CRC64;

Query Match 51.2%; Score 41.5; DB 1; Length 460;
 Best Local Similarity 81.8%; Pred. No. 61;
 Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 6 FP-PPPQETVT 15
 || ||| ||||
 Db 227 FPQPPPAETVT 237

RESULT 15

YK69_CAEEL

ID YK69_CAEEL STANDARD; PRT; 208 AA.
 AC P34347;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C29E4.9 in chromosome III.
 GN C29E4.9.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,

RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
 RA Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL; L23651; AAA27956.2; -.
 DR PIR; S44765; S44765.
 DR WormPep; C29E4.9; CE29199.
 KW Hypothetical protein.
 SQ SEQUENCE 208 AA; 23729 MW; FD4A57F6609EEB32 CRC64;

Query Match 50.6%; Score 41; DB 1; Length 208;
 Best Local Similarity 54.5%; Pred. No. 31;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQ 11
 : :: |||||
 Db 3 LFDLIFPPPPQ 13

Search completed: August 24, 2004, 15:43:35
 Job time : 10.0597 secs